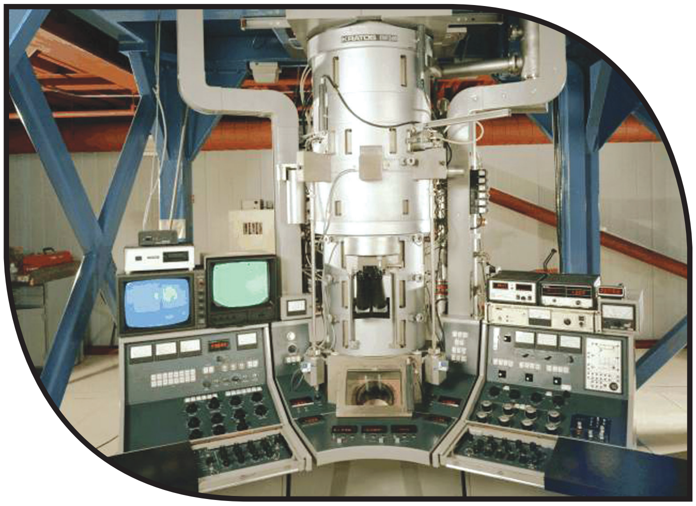
MODERN
PHYSICS

For B.Sc. / M.Sc. Students as per Syllabi of Different Indian Universities and also as per UGC Model Curriculum

R. MURUGESHAN
M.Sc., M.Phil.
Head of the Department of Physics (Retired)
Vivekananda College
Tiruvedakam West (Post)
Madurai District–625 217
TAMIL NADU

Er. KIRUTHIGA SIVAPRASATH
B.E.



S. CHAND & COMPANY PVT. LTD.
(AN ISO 9001 : 2008 COMPANY)
RAM NAGAR, NEW DELHI–110055

S. CHAND

S. CHAND & COMPANY PVT. LTD.
(An ISO 9001 : 2008 Company)
Head Office: 7361, RAM NAGAR, NEW DELHI - 110 055
Phone: 23672080-81-82, 9899107446, 9911310888
Fax: 91-11-23677446
Shop at: **schandgroup.com**; e-mail: **info@schandgroup.com**

Branches :
AHMEDABAD : 1st Floor, Heritage, Near Gujarat Vidhyapeeth, Ashram Road, **Ahmedabad** - 380 014,
Ph: 27541965, 27542369, ahmedabad@schandgroup.com
BENGALURU : No. 6, Ahuja Chambers, 1st Cross, Kumara Krupa Road, **Bengaluru** - 560 001,
Ph: 22268048, 22354008, bangalore@schandgroup.com
BHOPAL : Bajaj Tower, Plot No. 2&3, Lala Lajpat Rai Colony, Raisen Road, **Bhopal** - 462 011,
Ph: 4274723, 4209587. bhopal@schandgroup.com
CHANDIGARH : S.C.O. 2419-20, First Floor, Sector - 22-C (Near Aroma Hotel), **Chandigarh** -160 022,
Ph: 2725443, 2725446, chandigarh@schandgroup.com
CHENNAI : No.1, Whites Road, Opposite Express Avenue, Royapettah, **Chennai** - 600014
Ph. 28410027, 28410058, chennai@schandgroup.com
COIMBATORE : 1790, Trichy Road, LGB Colony, Ramanathapuram, **Coimbatore** - 6410045,
Ph: 2323620, 4217136 coimbatore@schandgroup.com **(Marketing Office)**
CUTTACK : 1st Floor, Bhartia Tower, Badambadi, **Cuttack** - 753 009, Ph: 2332580; 2332581,
cuttack@schandgroup.com
DEHRADUN : 1st Floor, 20, New Road, Near Dwarka Store, **Dehradun** - 248 001,
Ph: 2711101, 2710861, dehradun@schandgroup.com
GUWAHATI : Dilip Commercial (Ist floor), M.N. Road, Pan Bazar, **Guwahati** - 781 001,
Ph: 2738811, 2735640 guwahati@schandgroup.com
HYDERABAD : Padma Plaza, H.No. 3-4-630, Opp. Ratna College, Narayanaguda, **Hyderabad** - 500 029,
Ph: 27550194, 27550195, hyderabad@schandgroup.com
JAIPUR : 1st Floor, Nand Plaza, Hawa Sadak, Ajmer Road, **Jaipur** - 302 006,
Ph: 2219175, 2219176, jaipur@schandgroup.com
JALANDHAR : Mai Hiran Gate, **Jalandhar** - 144 008, Ph: 2401630, 5000630, jalandhar@schandgroup.com
KOCHI : Kachapilly Square, Mullassery Canal Road, Ernakulam, **Kochi** - 682 011,
Ph: 2378740, 2378207-08, cochin@schandgroup.com
KOLKATA : 285/J, Bipin Bihari Ganguli Street, **Kolkata** - 700 012, Ph: 22367459, 22373914,
kolkata@schandgroup.com
LUCKNOW : Mahabeer Market, 25 Gwynne Road, Aminabad, **Lucknow** - 226 018, Ph: 4076971, 4026791,
4065646, 4027188, lucknow@schandgroup.com
MUMBAI : Blackie House, IInd Floor, 103/5, Walchand Hirachand Marg, Opp. G.P.O., **Mumbai** - 400 001,
Ph: 22690881, 22610885, mumbai@schandgroup.com
NAGPUR : Karnal Bagh, Near Model Mill Chowk, **Nagpur** - 440 032, Ph: 2720523, 2777666
nagpur@schandgroup.com
PATNA : 104, Citicentre Ashok, Mahima Palace , Govind Mitra Road, **Patna** - 800 004, Ph: 2300489,
2302100, patna@schandgroup.com
PUNE : 291, Flat No.-16, Ganesh Gayatri Complex, IInd Floor, Somwarpeth, Near Jain Mandir,
Pune - 411 011, Ph: 64017298, pune@schandgroup.com **(Marketing Office)**
RAIPUR : Kailash Residency, Plot No. 4B, Bottle House Road, Shankar Nagar, **Raipur** - 492 007,
Ph: 2443142,Mb. : 09981200834, raipur@schandgroup.com **(Marketing Office)**
RANCHI : Flat No. 104, Sri Draupadi Smriti Apartments, (Near of Jaipal Singh Stadium) Neel Ratan Street,
Upper Bazar, **Ranchi** - 834 001, Ph: 2208761, ranchi@schandgroup.com **(Marketing Office)**
SILIGURI : 122, Raja Ram Mohan Roy Road, East Vivekanandapally, P.O., Siliguri, **Siliguri**-734001,
Dist., Jalpaiguri, (W.B.) Ph. 0353-2520750 **(Marketing Office)** siliguri@schandgroup.com
VISAKHAPATNAM : No. 49-54-15/53/8, Plot No. 7, 1st Floor, Opp. Radhakrishna Towers,
Seethammadhara North Extn., **Visakhapatnam** - 530 013, Ph-2782609 (M) 09440100555,
visakhapatnam@schandgroup.com **(Marketing Office)**

© 1984, R. Murugeshan & Er. Kiruthiga Sivaprasath

All rights reserved. No part of this publication may be reproduced or copied in any material form (including photo copying or storing it in any medium in form of graphics, electronic or mechanical means and whether or not transient or incidental to some other use of this publication) without written permission of the copyright owner. Any breach of this will entail legal action and prosecution without further notice.
***Jurisdiction** : All desputes with respect to this publication shall be subject to the jurisdiction of the Courts, tribunals and forums of New Delhi, India only.*

First Edition 1984
Subsequent Editions and Reprints 1994, 95 (Twice), 96, 98, 99, 2001, 2002, 2003, 2005, 2007, 2009, 2010, 2012, 2013
Seventeenth Revised Edition 2014

ISBN : 81-219-2801-X Code : 16D 333

PRINTED IN INDIA
By Rajendra Ravindra Printers Pvt. Ltd., 7361, Ram Nagar, New Delhi -110 055
and published by S. Chand & Company Pvt. Ltd., 7361, Ram Nagar, New Delhi -110 055.

Preface to the Seventeenth Edition

The present **Multicolour Edition has been Revised** by taking into account the 2010 Question Papers of various Indian Universities. Multicolour pictures have been added to enhance the content value and to give the students an idea of what they will be dealing in reality, and to bridge the gap between the theory and practice.

We are thankful to the Management Team and the Editorial Department of S. Chand & Company Pvt. Ltd. for all help and support in the publication of this book.

Any errors, omissions and suggestions, for the improvement of this volume brought to our notice, will be thankfully acknowledged and incorporated in the next edition.

AUTHORS

Preface to the First Edition

The purpose of this book is to give an introductory account of the basic principles of atomic, nuclear and solid state physics. It is meant to serve as a textbook for the B.Sc. Physics students of Indian Universities. SI system of units has been used throughout the text. Great care has been taken in dealing with the subject with modern outlook. A large number of questions and problems have been given at the end of each chapter.

All available standard books on the subject have been freely consulted during the preparation of this book. I acknowledge gratefully my indebtedness to the authors and publishers of different books on the subject. I am grateful to my revered teachers Dr. S. Hariharan M.Sc., Ph.D. (Leeds), Prof. S.N. Neelakantan Nambisan B.Sc. (Hons.), Prof. P. Radhakrishna Menon M.A., Prof. P. Balakrishnan M.A., Prof. K. Padmanabha Sharma M.Sc., and others of Govt. Victoria College, Palghat, Kerala, who contributed to my education in Modern Physics.

I am extremely thankful to Prof. K. Ramamurthi M. Com., Secretary and Prof. K. Subrahmanyam M.A. Principal, of our College for their encouragement and help.

I wish to dedicate this book at the holy feet of Srimad Swami Chidbavanandaji, President of our college. To Swamiji, I owe more than words can describe for his lifelong devotion and guidance in unfolding my hidden potentialities and developing my personality. I can only hope that this book reflects something of his insistence on lucid and concise writing.

Lastly, I would like to thank my Publishers, M/s. S. Chand & Company Ltd.

I especially want to thank the Proprietor, Mr. S.L. Gupta and the Madras Branch Manager Mr. S. Venkatachalam.

Tiruvedakam **AUTHOR**

Disclaimer : While the authors of this book have made every effort to avoid any mistakes or omissions and have used their skill, expertise and knowledge to the best of their capacity to provide accurate and updated informations, the authors and S. Chand do not give any representation or warranty with respect to the accuracy or completeness of the contents of this publication and are selling this publication on the condition and understanding that they shall not be made liable in any manner whatsoever. S.Chand and the authors expressly disclaim all and any liability/responsibility to any person, whether a purchaser or reader of this publication or not, in respect of anything and everything forming part of the contents of this publication. S. Chand shall not be responsible for any errors, omissions or damages arising out of the use of the information contained in this publication. Further, the appearance of the personal name, location, place and incidence, if any; in the illustrations used herein is purely coincidental and work of imagination. Thus the same should in no manner be termed as defamatory to any individual.

UGC MODEL CURRICULUM
B. Sc. Physics (General)

COURSE 7 : RELATIVITY, QUANTUM MECHANICS, ATOMIC, MOLECULAR AND NUCLEAR PHYSICS

7.1 Relativity (15)

Reference systems, inertial frames, Galilean invariance and conservation laws, propagation of light, Michelson-Morley experiment; search for ether. (5)

Postulates of the special theory of relativity, Lorentz transformations, length contraction, time dilation, velocity addition theorem, variation of mass with velocity, mass-energy equivalence, particle with a zero rest mass. (10)

7.2 Quantum Mechanics (30)

Origin of the quantum theory: Failure of classical physics to explain the phenomena such as black-body spectrum, photoelectric effect, Ritz combination principle in spectra, stability of an atom. Planck's radiation law, Einstein's explanation of photoelectric effect, Bohr's quantization of angular momentum and its applications to hydrogen atom, limitations of Bohr's theory. (5)

Wave-particle duality and uncertainty principle: de Broglie's hypothesis for matter waves; the concept of wave and group velocities, evidence for diffraction and interference of 'particles', experimental demonstration of matter waves. (4)

Consequence of de Broglie's concepts; quantisation in hydrogen atom; energies of a particle in a box, wave packets, Heisenberg's uncertainty relation for *p and x, its extension to energy and time.* (3)

Consequence of the uncertainty relation: gamma ray microscope, diffraction at a slit, particle in a box, position of electron in a Bohr orbit. (3)

Quantum Mechanics: Schrodinger's equation. Postulatory basis of quantum mechanics; operators, expectation values, transition probabilities, applications to particle in a one-and three-dimensional boxes, harmonic oscillator, reflection at a step potential, transmission across a potential barrier. (7)

Hydrogen atom; natural occurrence of *n, l and m quantum numbers, the related physical quantities, comparison with Bohr's theory.* (4)

7.3 Atomic Physics (15)

Spectra of hydrogen, deuteron and alkali atoms, spectral terms, doublet fine structure, screening constants for alkali spectra for *s, p, d, and f states, selection rules.* (6)

Singlet and triplet fine structure in alkaline earth spectra, L-S and J-J couplings. (3)

Weak spectra: continuous X-ray spectrum and its dependence on voltage, Duane and Hunt's law, Characteristic X-rays. Moseley's law, doublet structure of X-ray spectra, X-ray absorption spectra. (3)

7.4 Molecular Physics (15)

Discrete set of electronic energies of molecules, quantisation of vibrational and rotational energies, determination of internuclear distance, pure rotational and rotation vibration spectra. Dissociation limit for the ground and other electronic states, transition rules for pure vibration and electronic vibration spectra. (7)

Raman effect, Stokes and anti-Stokes lines, complimentary character of Raman and infrared spectra, experimental arrangements for Raman spectroscopy.

Spectroscopic techniques: Sources of excitation, prism and grating spectrographs for visible, UV and IR, absorption spectroscopy, double beam instruments, different recording systems. (5)

7.5 Nuclear Physics (15)

Interaction of charged particles and neutrons with matter, working of nuclear detectors, G-M counter, proportional counter and scintillation counter, cloud chambers, spark chamber, emulsions. (5)

Structure of nulei, basic properties (*l*, m , *Q and binding energy), deuteron binding energy, p-p and n-p scattering and general concepts of nuclear forces. Beta decay, range of alpha particle, Geiger-Nuttal law, Gamow's explanation of beta decay, alpha decay and continuous and discrete spectra.* (5)

Nuclear reactions, channels, compound nucleus, direct reaction (concepts). (3)

Shell model; liquid drop model, fission and fusion (concepts), energy production in stars by p-p and carbon cycles (concepts). (2)

COURSE 8: SOLID STATE PHYSICS, SOLID STATE DEVICES AND ELECTRONICS

8.1 Solid State Physics (30)

Overview: Crystalline and glassy forms, liquid crystals, glass transition. (2)

Structure: Crystal structure; periodicity, lattices and bases, fundamental translation vectors, unit cell, Winger-Seitz cell, allowed rotations, lattice types, lattice planes, common crystal structures. (5)

Laue's theory of X-ray diffraction, Bragg's law, Laue patterns. (2)

Bonding: Potential between a pair of atoms; Lennard-Jones potential, concept of cohesive energy, covalent, Van der Walls, ionic and metallic crystals. (3)

Magnetism: Atomic magnetic moment, magnetic susceptibility, Dia, para, and Ferromagnetism, Ferromagnetic domains, Hysteresis. (3)

Thermal properties: Lattice vibrations, simple harmonic oscillator, second order expansion of Lennard-Jones potential about the minimum, vibrations of one dimensional monatomic chain under harmonic and nearest neighbour interaction approximation, concept of phonons, density of modes (1-D). Debye model: lattice specific heat, low temperature limit, extension (conceptual) to 3-D. (8)

Band structure: Electrons in periodic potential; nearly free electron model (qualitative), energy bands, energy gap, metals, insulators, semiconductors. (3)

Motion of electrons: Free electrons, conduction electrons, electron collisions, mean free path, conductivity and Ohm's law. Density of states, Fermi energy, Fermi velocity, Fermi-Dirac distribution. (4)

8.2 Solid State Devices (15)

Semiconductors: Intrinsic semiconductors, electrons and holes, Fermi level. Temperature dependence of electron and hole concentrations. Doping: impurity states, *n and p type semiconductors, conductivity, mobility, Hall effect. Hall coefficient. (7)*

Semiconductor devices: Metal-semiconductor junction, p-n junction, majority and minority carriers, diode, Zener and tunnel diodes, light emitting diode, transistor, solar cell. (8)

8.3 Electronics (45)

Power supply: Diode as a circuit element, load line concept, rectification, ripple factor, zener diode, voltage stabilization, IC voltage regulation, characteristics of a transistor in CB, CE and CC mode, graphical analysis of the CE configuration, low frequency equivalent circuits, h-parameters, bias stability, thermal-runaway. (15)

Field effect transistors: JFET volt-ampere curves, biasing JFET, ac operation of JFET, source follower, Depletion and enhancement mode, MOSFET, biasing MOSFET, FET as variable voltage resistor, digital MOSFET circuits. (15)

Small signal amplifiers: General principles of operation, classification, distortion, RC coupled amplifier, gain, frequency response, input and output impedance, multistage amplifiers, transformer coupled amplifiers, Equivalent circuits at low, medium and high frequencies; emitter follower, low frequency common-source and common-drain amplifier, noise in electronic circuits. (15)

CONTENTS

PART – I : RELATIVITY

1. Theory of Relativity 3—25

1.1 Introduction; 1.2 Frame of Reference; 1.3 Newtonian Relativity; 1.4 Galilean Transformation Equations; 1.5 The Ether Hypothesis; 1.6 The Michelson-Morley Experiment; 1.7 Special Theory of Relativity; 1.8 The Lorentz Transformation Equations; 1.9 Length Contraction; 1.10 Time Dilation; 1.11 Relativity of Simultaneity; 1.12 Addition of Velocities; 1.13 Variation of Mass with Velocity; 1.14 Mass Energy Equivalence; 1.15 Minkowski's Four Dimensional Space-Time Continuum; 1.16 The General Theory of Relativity; 1.17 Predictions of General Relativity.

2. The Photon 26—30

2.1 Particle-Wave Duality; 2.2 Photons and Gravity; 2.3 Gravitational Red Shift; 2.4 To Show that the Rest Mass of a Photon is Zero.

3. Theory of Relativity II 31—32

3.1 Space-Time Diagrams; 3.2 Geometrical Representation of Simultaneity, Contraction and Dilation.

PART – II : ATOMIC PHYSICS

4. The Electron and Band Theory of Solids 35—47

4.1 Determination of the Electronic Charge : Millikan's Oil-Drop Method; 4.2 The Free Electron Theory of Metals; 4.3 Expression for Electrical Conductivity; 4.4 Expression for Thermal Conductivity; 4.5 Electron Microscope; 4.6 Band Theory of Solids; 4.7 Classification of Solids on the Basis of Band Theory; 4.8 Optical Properties of Solids; 4.9 Energy Bands: Alternative Analysis.

5. Positive Rays and Particle Properties of Waves 48—64

5.1 Discovery; 5.2 Properties of Positive Rays; 5.3 Positive Ray Analysis — Thomson's Parabola Method; 5.4 Aston's Mass Spectrograph; 5.5 Bainbridge's Mass Spectrograph; 5.6 Dempster's Mass Spectrograph; 5.7 Mass Defect and Packing Fraction; 5.8 Polarisation of X-rays; 5.9 Scattering of X-rays (Thomson's Formula); 5.10 Determination of the Number of Electrons per Atom; 5.11 Failure of Classical Mechanics.

6. Structure of the Atom 65—121

6.1 Introduction; 6.2 Rutherford's Experiments on Scattering of α-particles; 6.3 Theory of α-particle Scattering; 6.4 Bohr Atom Model; 6.5 Effect of Nuclear Motion on Atomic Spectra; 6.6 Evidences in Favour of Bohr's Theory; 6.7 Correspondence Principle; 6.8 Critical Potentials; 6.9 Atomic Excitation; 6.10 Experimental Determination of Critical Potentials; 6.11 Sommerfeld's Relativistic Atom Model; 6.12 The Vector Atom Model; 6.13 Quantum Numbers Associated with the Vector Atom Model; 6.14 Coupling Schemes; 6.15 The Pauli Exclusion Principle; 6.16 The Periodic Classification of Elements;

6.17 Some Examples of Electron Configurations with their Modern Symbolic Representations; 6.18 Magnetic Dipole Moment due to Orbital Motion of the Electron; 6.19 Magnetic Dipole Moment due to Spin; 6.20 The Stern and Gerlach Experiment; 6.21 Spin-Orbit Coupling; 6.22 Optical Spectra; 6.23 Zeeman Effect; 6.24 Larmor's Theorem; 6.25 Quantum Mechanical Explanation of the Normal Zeeman Effect; 6.26 Anomalous Zeeman Effect; 6.27 Paschen-Back Effect; 6.28 Stark Effect.

7. X-Rays **122—150**

7.1 Introduction; 7.2 Production of X-rays; 7.3 Spacing between Three Dimensional Lattice Planes; 7.4 The Absorption of X-rays; 7.5 X-ray Absorption Edges; 7.6 Bragg's Law; 7.7 The Bragg X-ray Spectrometer; 7.8 The Powder Crystal Method; 7.9 (a) The Laue Method; 7.9(b) Rotating-Crystal Method; 7.10 Symmetry Operations; 7.11 X-ray Spectra; 7.12 Characteristic X-ray Spectrum; 7.13 Moseley's Law; 7.14 Compton Scattering; 7.15 X-ray Crystallography; 7.16 Elements of Symmetry; 7.17 Bravais lattices; 7.18 Miller indices; 7.19 Typical Grystal Structures.

8. The Photoelectric Effect **151—160**

8.1 Introduction; 8.2 Lenard's Method to Determine e/m for Photoelectrons; 8.3 Richardson and Compton Experiment; 8.4 Experimental Investigations on the Photoelectric Effect; 8.5 Einstein's Photoelectric Equation; 8.6 Photoelectric Cells.

9. Planck's Quantum Theory **161—163**

10. Electron **164—165**

10.1 Dunnington's Method for Determining e/m.

PART – III : QUANTUM MECHANICS

11. Wave Mechanics **169—214**

11.1 Introduction; 11.2 Expression for Group Velocity; 11.3 Experimental Study of Matter Waves; 11.4 Heisenberg's Uncertainty Principle; 11.5 Wave Mechanical Atom Model; 11.6 Mathematical Proof of Uncertainty Principle for One Dimensional Wave-packet; 11.7 Basic Postulates of Wave Mechanics; 11.8 Derivation of Time-dependent Form of Schrodinger Equation; 11.9 Properties of the Wave Function; 11.10 The Particle in a Box – Infinite Square Well Potential; 11.11 Potential Step; 11.12 The Barrier Penetration Problem; 11.13 Linear Harmonic Oscillator; 11.14 The Hydrogen Atom; 11.15 The Rigid Rotator.

12. Quantum Mechanics II **215—232**

12.1 Postulates of Quantum Mechanics; 12.2 Probability Current Density; 12.3 The Free Particle; 12.4 Rectangular Potential Well; 12.5 The Square Well in Three Dimensions; 12.6 Ehrenfest's Theorem; 12.7 Mathematical Proof of Uncertainty Principle for One Dimensional Wave-packet; 12.8 Wave Velocity (or Phase Velocity);

13. Quantum Mechanics III **233—241**

13.1 Normalized Wave Functions of the Hydrogen Atom; 13.2 Expression for the Energy of the Electron of the Hydrogen Atom in the Ground State; 13.3 Significance of Various Quantum Numbers; 13.4 Electron Probability Density; 13.5 Radiative Transitions; 13.6 Selection Rules.

14. Quantum Mechanics IV **242—255**

14.1 Finite Square Potential Well; 14.2 The Parity Operator; 14.3 System of Two Identical Particles; 14.4 Commutator Algebra; 14.5 Orbital Angular Momentum Operator.

15. Quantum Mechanics V **256—274**

15.1 Dirac Delta Function; 15.2 Properties of δ-Function; 15.3 Fourier Transforms; 15.4 Application of Fourier Transform to a Rectangular Function; 15.5 Form of Wave Function in terms of Definite Momentum; 15.6 Probability Density; 15.7 To Show that the Energy Eigen Values are Real; 15.8 Deuteron Problem; 15.9 Dirac's 'BRA' and 'Ket' Vectors : Dirac Notation; 15.10 Linear Harmonic Oscillator (Operator method); 15.11 Heisenberg's Uncertainty Relation Derived from Operators.; 15.12 Equation of Motion : Evolution of System with Time, Constants of the Motion; 15.13 Angular Momentum, Defition; 15.14 Eigen Values of J^2 and J_z ; 15.15 Angular Momentum Matrices for $j = {}^1/_2$; 15.16 Angular Momentum Matrices for $j = 1$; 15.17 Addition of Angular Momenta : Clebsch Gordan Coefficients; 15.18 Addition of Angular Momenta.

16. Quantum Mechanics VI **275—289**

16.1 The Helium Atom; 16.2 Hydrogen Molecule; 16.3 Quantum Mechanical Hamiltonian of the Hydrogen Atom; 16.4 The Harmonic Oscillator; 16.5 Comparison of Classical and Quantum Ideas; 16.6 Scattering Theory; 16.7 Definitions of Differential Scattering Cross-section and Total Scattering Cross-section; 16.8 The Scattering Amplitude; 16.9 Partial Wave Analysis ; 16.10 Transformation from Centre of Mass to Laboratory Frame of Reference; 16.11 Reduction of the Two-Body Problem into One-Body Problem.

17. Three Dimensional Harmonic Oscillator **290—292**

17.1 Three Dimensional Harmonic Oscillator—(Spherically Symmetric Case).

18. Identical Particles and Spin **293—300**

18.1 Identical Particles; 18.2 Symmetric and Antisymmetric Wave Functions; 18.3 Construction of Symmetric and Antisymmetric Wave Functions from Unsymmetrised Functions; 18.4 Distinguishability of Identical Particles; 18.5 The Pauli Exclusion Principle; 18.6 Spin; 18.7 Pauli's Spin Matrices for Electron; 18.8 Commutation Relations Satisfied by Three Components σ_x, σ_y and σ_z ; 18.9 Two Component Wave Functions.

PART – IV : MOLECULAR PHYSICS

19. Lasers, Molecular Spectra and Raman Effect **303—327**

19.1 Induced Absorption, Spontaneous Emission and Stimulated Emission; 19.2 The Ruby Laser; 19.3 Helium-neon Laser ; 19.4 Semiconductor Laser; 19.5 Properties of a Laser Beam; 19.6 Molecular Spectra ; 19.7 Theory of the Origin of Pure Rotational Spectrum of a Molecule; 19.8 Theory of the Origin of the Vibration-rotation Spectrum of a Molecule; 19.9 Electronic Spectra of Molecules; 19.10 Rayleigh's Scattering; 19.11 Discovery; 19.12 Experimental Study of Raman Effect; 19.13 Quantum Theory of Raman Effect; 19.14 Applications; 19.15 Laser Raman Spectroscopy; 19.16 Rabi's Molecular Beam Magnetic Resonance Method to Determine Nuclear Magnetic Moment; 19.17 Molecular Orbital Theory of Hydrogen Molecule Ion; 19.18 Heitler-London Theory of Hydrogen Molecule.

20. Fine Structure of H_α Line **328—329**

20.1 Fine Structure of H_α Line.

21. Electronic Spectra of Molecules **330—333**

21.1 Electronic Spectra of Molecules; 21.2 Fluorescene and Phosphorescence.

22. Atomic Spectra **334—336**

22.1 Fine Width of Spectral Line; 22.2 The Width of Special Lines.

23. Molecular Spectra **337—351**

23.1 Origin of Molecular Spectra; 23.2 Nature of Molecular Spectra; 23.3 Different Modes of Molecular Excitation; 23.4 Factors Affecting Line Width of Molecular Spectra; 23.5 Factors Affecting Intensity of Molecular Spectra; 23.6 Born-Oppenheimer Approximation; 23.7 Rotation of Linear System; 23.8 Rotational Spectra of Rigid Linear Molecules (See Section 19.7); 23.9 Non-Rigid Rotator; 23.10 The Energy of a Diatomic Molecule; 23.11 Vibrating Diatomic Molecule as a Harmonic Oscillator; 23.12 Infra-red Spectra; 23.13 Theory of the Origin of the Vibration-Rotation Spectrum of a Molecule; 23.14 The Anharmonic Oscillator; 23.15 Fundamental and Overtone Frequencies; 23.16 Calculation of Force Constant.

24. Spectroscopic Techniques **352—368**

24.1 Introduction; 24.2 Constant Deviation Spectrograph; 24.3 Recording the Spectrum; 24.4 Enhancement of Spectra: Computer Averaging; 24.5 Ultraviolet Spectroscopy; 24.6 Quartz Spectrograph for Near U.V. Region; 24.7 Littrow Spectrograph; 24.8 Concave Grating Vacuum Spectrograph; 24.9 Infrared Spectroscopy; 24.10 Wadsworth Prism-Mirror Spectrograph; 24.11 Grating spectrograph for Infra-red ; 24.12 Absorption Spectroscopy; 24.13 Block Diagram of an Absorption Spectrometer; 24.14 Principle of a Double-Beam Spectrophotometer; 24.15. Double-beam Infrared Spectrometer; 24.16 Complementary Character of Raman and Infra-red Spectra.

25. Spectra of Alkali Elements **369—376**

25.1 Introduction; 25.2 Different Series in Alkali Spectra: Main Features; 25.3 Energy Levels in Alkali Spectra and Quantum Defect; 25.4 Nonterminating and Penetrating Orbits; 25.5 Explanation of Salient Features of Alkali Spectra; 25.6 The Spectrum of Lithium and other Hydrogen-like Species; 25.7 The Spectrum of Helium and the Alkaline Earths.

26. X-Ray Spectra **377—381**

26.1 Fine Structure of X-ray Levels; 26.2 Spin-relativity Doublets (Regular Doublets); 26.3 The Regular-doublet Law; 26.4 Screening Doublets; 26.5 Irregular-doublet Law; 26.6 X-ray Absorption Spectrum; 26.7 Auger Effect.

PART – V : NUCLEAR PHYSICS

27. Introduction to the Nucleus **385—407**

27.1 Introduction 27.2 Classification of Nuclei; 27.3 General properties of Nucleus; 27.4 Binding Energy; 27.5 Nuclear Stability; 27.6 Theories of Nuclear Composition; 27.7 Nuclear Forces; 27.8 Meson Theory of Nuclear Forces; 27.9 Models of Nuclear Structure; 27.10 The Liquid Drop Model; 27.11 The Shell Model; 27.12 The Collective Model; 27.13 Determination of Nuclear Radius-Mirror Nucler Method; 27.14 Nuclear Fission; 27.15 Applications of Semi-Empirical Mass-Formula.

28. Nuclear Structure **408—414**

28.1 Measurement of Nuclear Radius; 28.2 Parity; 28.3 Statistics of Nuclei; 28.4 Properties of Ground State of Deuterium Nucleus; 28.5 Nuclear Models—The Shell Model; 28.6 Fermi Gas Model of the Nucleus.

29. Detectors of Nuclear Radiations **415—431**

29.1 Introduction; 29.2 Interaction between Energetic Particles and Matter; 29.3 Ionization Chamber; 29.4 Solid-state Detectors; 29.5 Proportional Counter; 29.6 Geiger-Muller Counter; 29.7 The Wilson Cloud Chamber; 29.8 Diffusion Cloud Chamber; 29.9 Bubble Chamber; 29.10 Spark Chamber; 29.11 Nuclear Emulsions; 29.12 The Scintillation Counters ; 29.13 Cerenkov Counter.

30. Particle Accelerators **432—445**

30.1 Introduction; 30.2 Van de Graaff generator; 30.3 The Linear Accelerator; 30.4 The Cyclotron; 30.5 The Synchrocyclotron; 30.6 The Betatron; 30.7 The Synchrotrons; 30.8 The Proton Synchrotron (Bevatron, Cosmotron).

31. Radioactivity **446—479**

31.1 Discovery of Radioactivity; 31.2 Natural Radioactivity; 31.3 Alpha, Beta and Gamma Rays; 31.4 Properties of Alpha Rays; 31.5 Properties of Beta Rays; 31.6 Properties of Gamma Rays; 31.7 Determination of E/M of Alpha Particles; 31.8 Determination of the Charge of Alpha Particles; 31.9 Velocity of Alpha Particles; 31.10 Range of α-particles; 31.11 Experimental Measurement of the Range of α-particles; 31.12 Alpha Particle Disintegration Energy; 31.13 Alpha Particle Spectra; 31.14 Theory of Alpha Decay; 31.15 Gamow's Theory of Alpha Decay; 31.16 The Nature of Beta Particles; 31.17 Determination of e/m of Beta Particles : Kaufmann's Experiment; 31.18 Bucherer's Experiment-Increase of Beta Particle Mass with Velocity; 31.19 Beta Ray Spectra; 31.20 Magnetic Spectrograph ; 31.21 Origin of the Line and Continuous Spectrum; 31.22 The Neutrino Theory of Beta Decay; 31.23 Gamma Ray Spectra; 31.24 Determination of the Wavelength of Gamma Rays; 31.25 Origin of γ-rays; 31.26 Nuclear Isomerism; 31.27 Internal Conversion; 31.28 Mossbaur Effect; 31.29 Soddy Fajan's Displacement Law; 31.30 Law of Radioactive Disintegration; 31.31 The Mean Life; 31.32 Measurement of Decay Constants; 31.33 Units of Radioactivity; 31.34 Law of Successive Disintegration; 31.35 Radioactive Dating : The Age of the Earth; 31.36 Biological Effects of Nuclear Radiations; 31.37 Non-Conservation of Parity in Beta Decay.

32. Radiation Hazards **480—484**

32.1 Radiation Hazards; 32.2 Radiation Levels for Safety; 32.3 Radiation Protection Methods; 32.4 Nuclear Disasters; 32.5 Nuclear Waste Disposal; 32.6 Radiation Dosimeters.

33. Nuclear Reactions **485—487**

33.1 Neutron Thermalization; 33.2 Centre-of-Mass Coordinate System.

34. Artificial Transmutation of Elements **488—509**

34.1 The Discovery of Artificial Transmutation; 34.2 Bohr's Theory of Nuclear Disintegration; 34.3 The Q-Value Equation for a Nuclear Reaction; 34.4 Nuclear Reactions; 34.5 Energy Balance in Nuclear Reactions and the Q-value; 34.6 Threshold Energy of an Endoergic Reaction; 34.7 Nuclear Transmutations; 34.8 The Scattering Cross-Section; 34.9 Discovery; 34.10 Preparation of Radioelements; 34.11 Applications of Radioisotopes;

34.12 The Discovery of the Neutron; 34.13 Basic Properties of the Neutron; 34.14 Classification of Neutrons; 34.15 Neutron Sources; 34.16 Neutron Detection; 34.17 Neutron Collimator.

35. Nuclear Fission and Fusion 510—523

35.1 Discovery ; 35.2 Nuclear Fission; 35.3 Energy Released in Fission; 35.4 Chain Reaction; 35.5 Atom Bomb; 35.6 Nuclear Reactors; 35.7 Nuclear Fusion; 35.8 Source of Stellar Energy; 35.9 Thermonuclear Reactions; 35.10 Transuranic Elements.

36. Nuclear Reactors 524—529

36.1 Pressurized Water Reactor (PWR); 36.2 Boiling Water Reactor (BWR); 36.3 Fast Breeder Reactor; 36.4 Neutron Cycle in a Thermal Nuclear Reactor; 36.5 Fusion Reactors; 36.6 Plasma Confinement.

37. Cosmic Rays and the Universe 530—539

37.1 Discovery of Cosmic Rays; 37.2 Latitude Effect; 37.3 The East-West Effect (Azimuth Effect); 37.4 Altitude Effect; 37.5 Primary Cosmic Rays; 37.6 Secondary Cosmic Rays; 37.7 Cosmic Ray Showers; 37.8 Discovery of Positron; 37.9 The Mesons; 37.10 Van Allen belts; 37.11 Origin of Cosmic Rays; 37.12 The Big Bang Theory; 37.13 Thermal History of the Universe; 37.14 Hubble's Law; 37.15 The Future of the Universe.

38. Elementary Particles I 540—551

38.1 Introduction; 38.2 Particles and Anti-particles; 38.3 Antimatter; 38.4 The Fundamental Interactions; 38.5 Elementary-particle quantum numbers; 38.6 Conservation Laws and Symmetry; 38.7 The Quark Model.

39. Elementary Particles II 552—554

39.1 Quantum Chromodynamics; 39.2 The Standard Model; 39.3 Unification of Interactions.

40. Interaction of Charged Particles with Matter 555—560

40.1 Rate of Loss of Energy of a Charged Particle Transversing a Material Medium; 40.2 Bremsstrahlung; 40.3 Neutron Decay; 40.4 Stability of Proton; 40.5 Resonance Particles; 40.6 Detection of Neutrino.

PART – VI : SOLID STATE PHYSICS

41. Solid State Physics I 563—579

41.1 Bonding in Crystals; 41.2 Ionic bond; 41.3 Covalent Bond; 41.4 Metallic Bond; 41.5 Molecular Bond; 41.6 Hydrogen Bond; 41.7 Calculation of Repulsive Exponent n; 41.8 Born-Haber Cycle; 41.9 Some Simple Crystal Structures; 41.10 Specific Heat Capacity of Solids; 41.11 Einstein's Theory of Specific Heat Capacity; 41.12 Debye's Theory of Specific Heat Capacity of a Solid; 41.13 Superconductivity; 41.14 The Meissner Effect; 41.15 The BCS Theory.

42. Solid State Physics II 580—588

42.1 Classification of Superconductors; 42.2 Principle of Magnetic Lavitation; 42.3 X-ray Diffraction: Laue Equations; 42.4 Magnetoresistance; 42.5 Rotating Crystal Method; 42.6 Powder Method (Debye-Scherer Method)

43. Liquid Crystals **589—595**

43.1 Liquid Crystals; 43.2 Thermotropic Liquid Crystals; 43.3 Lyotropic Liquid Crystals; 43.4 Applications; 43.5 Glass; 43.6 Glass Transition Temperature; 43.7 Metallic Glasses; 43.8 Quasicrystals.

44. Superconductivity I **596—603**

44.1 The Josephson Effect; 44.2 Theory of DC Josephson Effect; 44.3 Theory of AC Josephson Effect; 44.4 Flux Quantization; 44.5 High-T_c Superconductivity; 44.6 Applications.

45. Superconductivity II **604—606**

45.1 London Equation; 45.2 Energy Gap; 45.3 Microwave and Infrared Absorptions.

46. Reciprocal Lattice, Crystal Binding, Energy Bands **607—634**

46.1 Reciprocal Lattice; 46.2 Reciprocal Lattice Vectors; 46.3 Properties of Reciprocal Lattice; 46.4 Bragg's law for Reciprocal Lattice; 46.5 Brillouin Zones; 46.6 Reciprocal Lattice to Simple Cubic (sc) Lattice; 46.7 Reciprocal Lattice to bcc Lattice; 46.8 Reciprocal Lattice to fcc Lattice; 46.9 Atomic Scattering Factor; 46.10 The Geometrical Structure Factor; 46.11 Cohesive Energy of Inert Gas Crystals; 46.12 Binding Energy of Ionic Crystals; 46.13 Evaluation of the Madelung Constant; 46.14 Bloch Theorem; 46.15 Kronig-Penney Model; 46.16 Effective Mass of Electron; 46.17 Interatomic Potential; 46.18 Second Order Expansion of Potential Energy; 46.19 Second Order Expansion of Lennard-Jones Potential about the Minimum ; 46.20 Simple Harmonic Oscillator.

47. Magnetic Properties of Materials **635—642**

47.1 Introduction; 47.2 Langevin's Theory of Diamagnetism; 47.3 Langevin's Theory of Paramagnetism; 47.4 Ferromagnetism.

48. Magnetism **643—646**

48.1 Weiss Theory of Ferromagnetism; 48.2 Nuclear Magnetic Resonance.

49. Quantum Theory of Paramagnetism **647—651**

49.1 Quantum Theory of Paramagnetism; 49.2 Cooling by Adiabatic Demagnetization of a Paramagnetic Salt.

50. Lattice Vibrations and Phonons **652—663**

50.1 Introduction; 50.2 Vibrations of Monatomic Linear Lattice; 50.3 Vibrational Modes of a Finite One-dimensional Lattice of Identical Atoms; 50.4 Vibrational Modes of a Diatomic Linear Lattice ; 50.5 Quantisation of Lattice Vibrations; 50.6 Inelastic Scattering of Neutrons by Phonons.

51. Electrical Conductivity and Ohm's Law **664—673**

51.1 Electrical Conductivity; 51.2 Experimental Electrical Resistivity of Metals; 51.3 Conduction Electrons; 51.4 Electron Collisions; 51.5 Derivation of Boltzmann Transport Equation; 51.6 Paramagnetic Susceptibility of Conduction Electrons.

52. Free Electron Model of Metals **674—680**

52.1 Sommerfeld's Model; 52.2 Fermi-Energy; 52.3 Sommerfeld Theory of Electrical Conductivity.

53. Free Electron Fermi Gas **681—685**

53.1 Energy Levels in One Dimension; 53.2 Free Electron Gas in Three Dimensions.

54. Dielectric Properties of Solids **686—695**

54.1 Introduction; 54.2 Polarization; 54.3 Macroscopic Electric Field; 54.4 Depolarization Field; 54.5 Local Electric Field at an Atom; 54.6 Dielectric Constant; 54.7 Polarizability; 54.8 Derivation of the Clausius-Mossotti Relation; 54.9 Frequency Dependence of total Polarizability; 54.10 Electronic Polarizability; 54.11 Classical Theory of Electronic Polarizability.

PART – VII : SOLID STATE DEVICES

55. Semiconductors **699—726**

55.1 Electron and Hole Concentrations in Intrinsic Semiconductor in Thermal Equilibrium; 55.2 Fermi Level in N-type Extrinsic Semiconductor; 55.3 Fermi Level in P-type Extrinsic Semiconductor; 55.4 Electrical Conductivity; 55.5 Diffusion Current; 55.6 Total Current; 55.7 Carrier Generation and Recombination : Carrier Life Times; 55.8 The Continuity Equation; 55.9 Derivation of Expressions for the Width of Depletion Layer and Height of Potential Barrier; 55.10 Derivation of Voltage Current (V-I) Equation for a P-N Junction Diode; 55.11 P-N Diode Clipping Circuits; 55.12 P-N Diode Clamping Circuits; 55.13 Differentiating Circuit; 55.14 Integrating Circuit; 55.15 Metal-Semiconductor Junction; 55.16 Schottky Diode.

56. Optoelectronic Devices **727—735**

56.1 Photo Diode; 56.2 Photo Transistor; 56.3 Photoconductors; 56.4 Light Emitting Diode (LED); 56.5 Liquid Crystal Display (LCD); 56.6 The Solar Cell.

57. Hall Effect **736—738**

57.1 Hall Effect; 57.2 Experimental Determination of Hall Coefficient.

PART – VIII : ELECTRONICS

58. Electronics **741—781**

58.1 Formation of p-n Junction Diode; 58.2 Forward and Reverse Biasing of a Junction Diode; 58.3 V-I Characteristics of a Junction Diode; 58.4 The Zener Diode; 58.5 Experiment to Study the Characteristics of the Zener Diode; 58.6 Zener Diode as Voltage Regulator; 58.7 P-N Junction Diode as Rectifier; 58.8 Diode as a Half-Wave Rectifier; 58.9 Mathematical Analysis; 58.10 Full-Wave Rectifier; 58.11 Mathematical Analysis; 58.12 Bridge Rectifier; 58.13 Filter Circuits; 58.14 Junction Transistor; 58.15 Common Emitter Characteristics of a Transistor; 58.16 Common-Emitter Transistor Amplifier; 58.17 Practical Circuit of Common Emitter Amplifier; 58.18 A.C. Equivalent Circuit; 58.19 Common-Base Transistor Amplifier; 58.20 Oscillators; 58.21 Condition for Self-excitation (Barkhausen Criterion for Oscillations); 58.22 Hartley Oscillator; 58.23 Colpitt's Oscillator; 58.24 Tuned Collector Oscillator; 58.25 Tunnel Diode; 58.26 Field-Effect Transistors; 58.27 Drain Characteristics of an *n* Channel JFET; 58.28 Principles of n-Channel FET Operation; 58.29 Integrated Circuit (IC); 58.30 Fabrication of a Monolithic IC; 58.31 Modulation and Detection; 58.32 Amplitude Modulation; 58.33 Power in an Amplitude Modulated Wave; 58.34 Frequency Modulation; 58.35 Demodulation or Detection; 58.36 Block Diagram of AM Transmitting System; 58.37 Block Diagram of AM

Recievers—Superheterodyne Reciever; 58.38 Characteristics of a Radio-receiver; 58.39 Block Diagram of TV Transmitting System; 58.40 Block Diagram of TV Receiving System.

59. Transistor Biasing **782—789**

59.1 Transistor Biasing; 59.2 Fixed Bias or Base Bias; 59.3 Collector to Base Bias Circuit; 59.4 Self Bias or Emitter Bias (Voltage-Divider Bias); 59.5 Thermal Runaway; 59.6 Characteristics of a Transistor Connected in Common Base Configuration; 59.7 C – C Configuration.

60. Electronics II **790—803**

60.1 The D.C. Load Line; 60.2 A. C. Load Line; 60.3 Satellite Communication; 60.4 Television Camera — The Image Orthicon; 60.5 Scanning Process in TV; 60.6 Colour Television; 60.7 Colour Television Camera; 60.8 Transmission of Colour TV Signal; 60.9 Reception of Colour TV Signal; 60.10 Synchronization in TV; 60.11 Television Channel Width; 60.12 SCR Applications; Silicon-Controlled Rectifier.

61. Hybrid Parameters **804—807**

61.1 Two-port Representation of a Transistor; 61.2 The Hybrid Equivalent Circuit ; 61.3 Determination of the *h* Parameters.

62. Transistor Amplifiers **808—818**

62.1 Hybrid Equivalent Circuit of Common Emitter Amplifier; 62.2 Analysis of a Transistor CE Amplifier Using *h*-Parameters; 62.3 Power Amplifier; 62.4 Class A Power Amplifier; 62.5 Transformer Coupled Class A Power Amplifier; 62.6 Comparison of Amplifier Configurations; 62.7 Class B Push-Pull Amplifier.

63. Small Signal Amplifiers **819—828**

63.1 Resistance-Capacitance (R-C) Coupled Amplifiers; 63.2 Transformer Coupled (TC) Amplifier; 63.3 Emitter Follower; 63.4 Noise in Electronic Circuits; 63.5 Distortion in Amplifiers.

64. FET Amplifier **829—834**

64.1 Common Source FET Amplifier; 64.2 Advantages of JFET over BJT; 64.3 The Equivalent Circuit for the FET (Low frequency model); 64.4 JFET Source-Follower (Common-Drain) Amplifier; 64.5 FET as Voltage-Variable Resistor (VVR).

65. Operational Amplifiers **835—844**

65.1 Introduction; 65.2 Characteristics of an OP AMP; 65.3 The Common Mode Rejection Ratio (CMRR); 65.4 The Slew Rate; 65.5 Inverting Amplifier; 65.6 Adder or Summing Amplifier; 65.7 Op-Amp as Integrator; 65.8 Op-Amp as Differentiator; 65.9 Difference Amplifier or Subtractor; 65.10 Logarithmic Amplifier; 65.11 Op-Amp as Comparator; 65.12 Non-inverting Operational Amplifier; 65.13 Voltage Follower.

66. Oscillators **845—859**

66.1 Hartley Oscillator; 66.2 Colpitt's Oscillator; 66.3 Wien-Bridge Oscillator; 66.4 Phase Shift Oscillator; 66.5 Crystal Oscillator; 66.6 Wien Bridge Oscillator.

67. Multivibrators and Logic Gates **860—872**

67.1 Astable Multivibrator; 67.2 Monostable Multivibrator; 67.3 Bistable Multivibrator or Flip-Flop Circuit; 67.4 The RS Flip-Flop; 67.5 The J–K Flip-Flop; 67.6 The NOR Gate; 67.7 NOR Gate is a Universal Gate; 67.8 The NAND Gate; 67.9 NAND Gate is Universal Gate; 67.10 The EXCLUSIVE OR gate.

68. Metal Oxide Semiconductor Field Effect Transistor (MOSFET) **873—878**

68.1 Introduction; 68.2 The Depletion MOSFET; 68.3 Static Characteristics of Depletion MOSFET; 68.4 The Enhancement MOSFET; 68.5 Characteristics of Enhancement MOSFET; 68.6 MOSFET Biasing.

69. IC Voltage Regulators **879—881**

69.1 Introduction; 69.2 Fixed Voltage Regulators; 69.3 Regulated Dual Supplies; 69.4 Adjustable Voltage Regulator.

70. Digital Electronics **882—900**

70.1 Decimal Number System; 70.2 Binary Number System; 70.3 Conversion of Binary Number into Decimal Number; 70.4 Conversion of Decimal Number into Binary Number; 70.5 Binary Addition; 70.6 Binary Subtraction; 70.7 Complement Method of Subtraction; 70.8 Binary Multiplication; 70.9 Binary Division; 70.10 Octal Number System; 70.11 Hexadecimal Number System; 70.12 Boolean Algebra; 70.13 Boolean Addition and Multiplication; 70.14 Logic Expressions; 70.15 Commutative Laws; 70.16 Associative Laws; 70.17 Distributive Law; 70.18 Rules for Boolean Algebra; 70.19 De Morgan's Theorems; 70.20 Digital Logic Gates; 70.21 The OR Gate; 70.22 The AND Gate; 70.23 The NOT Gate; 70.24 Flip-Flop; 70.25 Master-Slave Flip-Flop; 70.26 Master-Slave JK Flip-Flop; 70.27 Counters; 70.28 The 7493 A Four-Bit Binary Counter; 70.29 Shift Register; 70.30 Serial in-Serial out Shift Register.

71. Half Adder and Full Adder **901—905**

71.1 Half-Adder; 71.2 Full Adder; 71.3 Parallel Binary Adder.

72. BJT Transistor Modeling **906—914**

72.1 Introduction; 72.2 The Important Parameters : Z_i, Z_o, A_v, A_i; 72.3 The r_e Transister Model.

73. Unijunction Transistor (UJT) **915—919**

73.1 Construction; 73.2 Equivalent Circuit of a UJT; 73.3 Operation of UJT; 73.4 Static Emitter-Characteristic Curve for a UJT; 73.5 Applications of UJT.

74. Feedback Amplifiers **920—924**

74.1 Feedback Principle; 74.2 The Effect of Feedback on the Gain of an Amplifier; 74.3 Advantages of Negative Feedback; 74.4 Practical Feedback Circuits.

PART – IX : STATISTICAL MECHANICS

75. Statistical Mechanics I **927—937**

75.1 Introduction; 75.2 Phase Space; 75.3 Maxwell-Boltzmann Distribution Law; 75.4 Molecular Energies in an Ideal Gas; 75.5 Bose-Einstein Distribution Law; 75.6 Fermi-Dirac Distribution Law; 75.7 Comparison of the three Distribution Laws; 75.8 Black-Body Radiation; 75.9 Rayleigh-Jeans Formula; 75.10 Planck Radiation Formula; 75.11 Wien's Displacement Law; 75.12 Stefan-Boltzmann Law from Planck's formula.

76. Statistical Mechanics II **938—943**

76.1 Macroscopic and Microscopic Descriptions; 76.2 Ensembles; 76.3 Probability; 76.4 Thermodynamic Probability; 76.5 Boltzmann's Theorem on Entropy and Probability; 76.6 Fundamental Postulates of Statistical Mechanics; 76.7 Statistical Equilibrium; 76.8 Quantum Statistics; 76.9 Electron Gas.

PART – X : MATERIAL SCIENCE

77. Material Science 947—1001

77.1 The Seven Crystal Systems; 77.2 Absence of Five-fold Symmetry Axis; 77.3 Definitions; 77.4 Simple Cubic (SC) Structure ; 77.5 Body Centred Cubic (BCC) Structure; 77.6 Face Centred Cubic (FCC) Structure; 77.7 Calculation of Lattice Constant; 77.8 Hexagonal Close Packed (HCP) Structure; 77.9 Diamond Cubic Structure; 77.10 Summary; 77.11 Defects in Crystals; 77.12 Classification of Crystal Imperfections (or Defects) 77.13 Point Defects; 77.14 To Derive an Expression for the Number of Schottky Defects at a given Temperature ; 77.15 To Derive an Expression for the Number of Frenkel Defects at a given Temperature; 77.16 To Derive an Expression for the Number of Vacancies at a given Temperature; 77.17 Line Defects; 77.18 Surface Defects; 77.19 Effects of Crystal Imperfection; 77.20 Ferrites and their Application; 77.21 Ferromagnetic Domains; 77.22 Four Types of Energy Involved in the Process of Domain Growth; 77.23 Reversible and Irreversible Domains; 77.24 Explanation of Hysteresis Curve on the Basis of Domain Theory; 77.25 Crystal Growth; 77.26 Crystal Growth from Solution; 77.27 Crystal Growth from Melt-Czochralski Method; 77.28 Crystal Growth from Vapour Phase; 77.29 Weiss Theory of Paramagnetism; 77.30 Heisenberg's Interpretation of Internal Field or Weiss Molecular Field; 77.31 Quantum Theory of Ferromagnetism; 77.32 Fundamental Definitions in Dieletrics; 77.33 Different Types of Electric Polarisation; 77.34 Total Polarisation; 77.35 Effect of Temperature on Polarisation; 77.36 Dielectric Loss; 77.37 Determination of Dielectric Constant of a Dielectric Material; 77.38 Dielectric Breakdown; 77.39 Properties and Applications of Different Types of Insulating Materials.

PART – XI : ASTROPHYSICS

78. Astrophysics 1005—1018

78.1 Introduction; 78.2 Classification of Stars – The Harvard Classification System; 78.3 Hertzsprung — Russel Diagram; 78.4 Luminosity of a Star; 78.5 Stellar Evolution; 78.6 White Dwarfs; 78.7 Electrons in a White Dwarf Star; 78.8 Chandrasekhar Limit; 78.9 Neutron Stars; 78.10 Black Holes; 78.11 Supernova Explosion; 78.12 Photon Diffusion Time; 78.13 Gravitational Potential Energy of a Star; 78.14 Internal Temperature of a Star; 78.15 Internal Pressure of a Star.

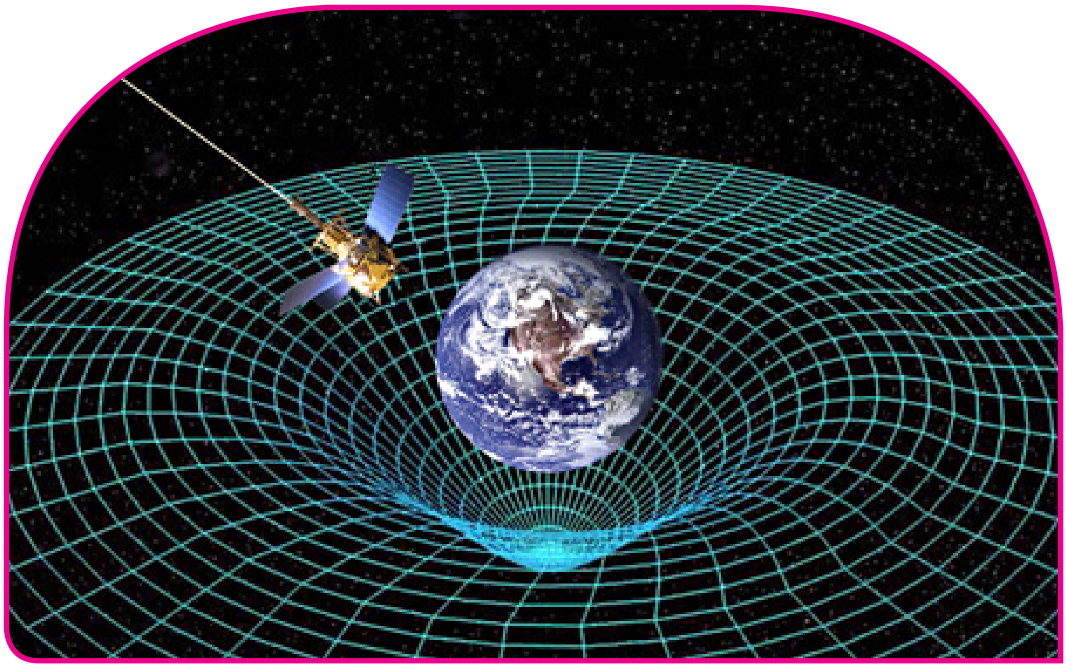
79. Nanotechnology 1019—1040

79.1 Introduction; 79.2 Nanoparticles; 79.3 Metal Nanoclusters; 79.4 Semiconducting Nanoparticles; 79.5 Rare Gas and Molecular Clusters; 79.6 Carbon Clusters; 79.7 Carbon Nanotubes; 79.8 Applications of Carbon Nanotubes; 79.9 Quantum Nanostructures; 79.10 Preparation of Quantum Nanostructures; 79.11 Size and Dimensionality Effects; 79.12 Applications of Quantum Nanostructures; 79.13 Photolithography; 79.14 Scanning Tunnelling Microscopy (STM); 79.15 Atomic Force Microscopy (AFM).

PART–I

RELATIVITY

1. Theory of Relativity
2. The Photon
3. Theory of Relativity II



CHAPTER

1

THEORY OF RELATIVITY

AT A GLANCE

1.1	Introduction	**1.2**	Frame of Reference
1.3	Newtonian Relativity	**1.4**	Galilean Transformation Equations
1.5	The Ether Hypothesis	**1.6**	The Michelson-Morley Experiment
1.7	Special Theory of Relativity	**1.8**	The Lorentz Transformation Equations
1.9	Length Contraction	**1.10**	Time Dilation
1.11	Relativity of Simultaneity	**1.12**	Addition of Velocities
1.13	Variation of Mass with Velocity	**1.14**	Mass Energy Equivalence
1.15	Minkowski's Four Dimensional Space-Time Continuum		
1.16	The General Theory of Relativity	**1.17**	Predictions of General Relativity

1.1 Introduction

Classical or Newtonian mechanics deals with the motions of bodies travelling at velocities that are very much less than the velocity of light. According to it, the three fundamental concepts of Physics, *viz.*, space, time and mass are all absolute and invariant.

Concept of space. Newton assumed that space is absolute and "exists in itself, without relation to anything external and remains unaffected under all circumstances". This means to say that the length of an object is independent of the conditions under which it is measured, such as the motion of the object or the experimenter.

Concept of time. According to Newton, time is absolute 'by its very nature flowing uniformly without reference to anything external'. Hence there is a universal time flowing at a constant rate, unaffected by the motion or position of objects and observers. This implies two things:

(1) The interval of time between two events has the same value for all observers, irrespective of their state of motion.

(2) If two events are simultaneous for an observer, they are simultaneous for all observers, irrespective of their state of motion, *i.e.*, *simultaneity is absolute.*

Concept of mass. In Newtonian mechanics,

(1) The mass of a body does not depend on the velocity of its motion.

(2) The mass of an isolated system of bodies does not change with any processes occurring within the system (*law of conservation of mass*).

1.2 Frame of Reference

A system of co-ordinate axes which defines the position of a particle in two or three dimensional space is called a *frame of reference*. The simplest frame of reference is the familiar Cartesian system of co-ordinates, in which the position of the particle is specified by its three co-ordinates *x,y,z,* along the three perpendicular axes. In Fig. 1.1 we have indicated two observers *O* and *O'* and a particle *P*. These observers use frames of reference *XYZ* and *X' Y' Z',* respectively. If *O* and *O'* are at rest, they will observe the same motion of *P*. But if *O* and *O'* are in relative motion, their observation of the motion of *P* will be different.

Unaccelerated reference frames in uniform motion of translation relative to one another are called *Galilean frames or inertial frames.*

Accelerated frames are called *non-inertial* frames.

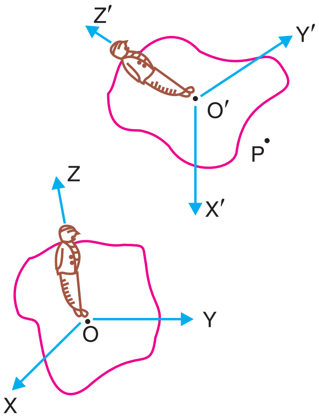


Fig. 1.1

1.3 Newtonian Relativity

The Newtonian principle of relativity may be stated as "*Absolute motion, which is the translation of a body from one absolute place to another absolute place, can never be detected. Translatory motion can be perceived only in the form of motion relative to other material bodies*".

This implies that if we are drifting along at a uniform speed in a closed spaceship, all the phenomena observed and all the experiments performed inside the ship will appear to be the same as if the ship were not in motion. This means that the fundamental physical laws and principles are identical in all inertial frames of reference. This is the concept of Newtonian relativity.

1.4 Galilean Transformation Equations

Let *S* and *S′* be two inertial frames (Fig. 1.2).

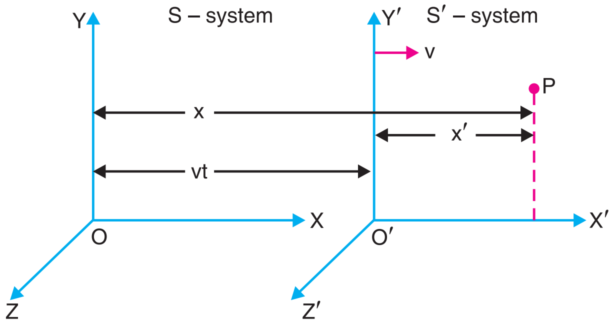


Fig. 1.2

Let S be at rest and S' move with uniform velocity v along the positive X direction. We assume that $v << c$. Let the origins of the two frames coincide at $t = 0$. Suppose some event occurs at the point P. The observer O in frame S determines the position of the event by the coordinates x, y, z. The observer O' in frame S' determines the position of the event by the coordinates x', y', z'. There is no relative motion between S and S' along the axes of Y and Z. Hence we have $y = y'$ and $z = z'$. Let the time proceed at the same rate in both frames.

The distance moved by S' in the positive X-direction in time $t = vt$. So the X coordinates of the two frames differ by vt. Hence, $x' = x - vt$. Then the transformation equations from S to S' are given by,

$$\mathbf{x' = x - vt \;...\; (1) \qquad y' = y \;...\; (2) \qquad z' = z \;...\; (3) \qquad t' = t \;...\; (4)}$$

Notes

(1) *The inverse transformation* (from S' to S) will be

$$x = x' + vt,\ y = y',\ z = z' \text{ and } t = t'.$$

(2) In general, the transformation of velocities from one to the other system is obtained by taking time derivatives. When the relative motion of the two frames is confined to the X-direction, the transformation becomes,

$$\frac{dx'}{dt} = \frac{dx}{dt} - v \quad i.e., \quad u' = u - v;\ \frac{dy'}{dt} = \frac{dy}{dt};\ \frac{dz'}{dt} = \frac{dz}{dt}.$$

(3) Let a and a' be the accelerations of the particle in S and S'. We have $a = \frac{du}{dt}$ and $a' = \frac{du'}{dt}$.

We have seen above that $u' = u - v$.

$$\therefore \qquad \frac{du'}{dt} = \frac{du}{dt} \qquad \text{(since } v \text{ is constant)}$$

or $a' = a$. *i.e., the accelerations, as measured by the two observers in the two frames, are the same.* Hence we say that *acceleration is invariant under Galilean transformation.*

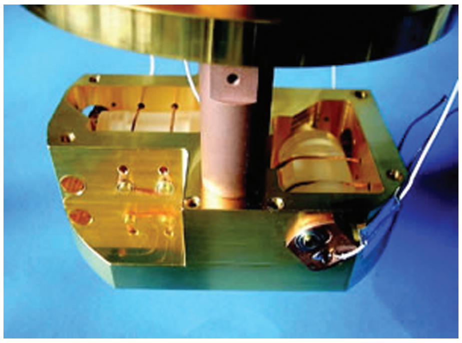
1.5 The Ether Hypothesis

A material medium is a necessity for the propagation of waves. It was considered that light propagates through *ether* as the sound waves propagate through air. Ether pervades all space. An interesting question is whether relative motion between the earth and ether can be detected. If such a motion can be detected, we can choose a fixed frame of reference in a stationary ether. Then we can express all motion relative to this frame. In 1887, Michelson and Morley, set out to measure the relative velocity of earth with respect to the ether. The principle of the experiment lies in noting the shift in fringes in the Michelson interferometer due to the difference in time taken by light to travel along and opposite the direction of motion of the earth. The time taken by a beam of light to travel along the direction of motion of earth is greater than that to travel distance opposite to the direction of motion of the earth. Surprisingly, despite best efforts, the presence of ether could not be detected.

1.6 The Michelson-Morley Experiment

A beam of light from a monochromatic source S falls on a half-silvered glass plate P, placed at an angle of 45° to the beam. The incident beam is split up into two parts by P (Fig. 1.3). The reflected portion travels in a direction at right angles to the incident beam, falls normally at B on the plane mirror M_1 and is reflected back to P. It gets refracted through P and enters the telescope T. The transmitted portion travels along the direction of the initial beam, falls normally on mirror M_2 at A and is reflected back to P. After reflection from the back surface of P, it enters the telescope T. The two reflected beams interfere and the interference fringes are viewed with the help of the telescope

T. The beam reflected upwards to M_1 traverses the thickness of plate P thrice whereas the beam refracted on to mirror M_2 traverses P only once. The effective distance of the mirrors M_1 and M_2 from the plate P is made to be the same by the use of a compensating plate not shown in figure.



Michelson-Morley Experimental Setup.

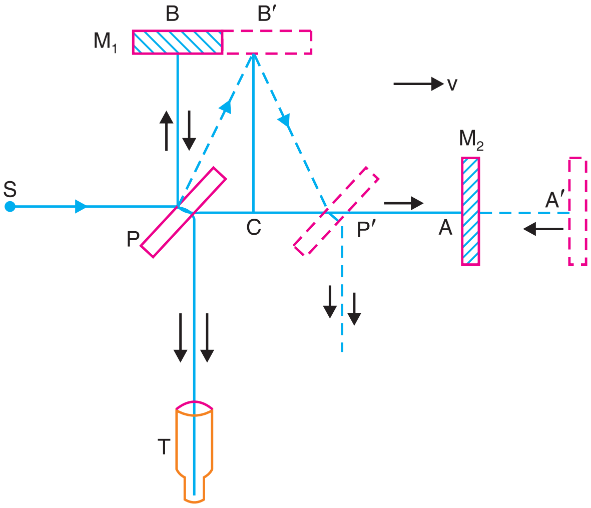


Fig. 1.3

The whole apparatus was floating on mercury. One arm (PA) was pointed in the direction of earth's motion round the sun and the other (PB) was pointed perpendicular to this motion. The paths of the two beams and the positions of their reflections from M_1 and M_2 will be as shown by the dotted lines.

Assume that the velocity of the apparatus (or earth) relative to fixed ether is v in the direction PA. The relative velocity of a light ray travelling along PA is $(c - v)$ while its value would be $(c + v)$ for the returning ray. Let $PA = PB = d$.

Time taken by light to travel from P to $A = d/(c - v)$.

Time taken by light to travel from A to $P = d/(c + v)$.

$\therefore$ Total time taken by light to travel from P to A and back

$$t = \frac{d}{c-v} + \frac{d}{c+v} = \frac{2cd}{c^2-v^2} = \frac{2d}{c}\left(1+\frac{v^2}{c^2}\right) \qquad \text{...(1)}$$

Now, consider the ray moving upwards from P to B. It will strike the mirror M_1 not at B but at B' due to the motion of the earth. If t_1 is the time taken by the ray starting from P to reach M_1, then $PB' = ct_1$ and $BB' = vt_1$.

The total path of the ray until it returns to $P = PB'\,P'$.

Now $PB'\,P' = PB' + B'\,P' = 2PB'$, since $PB' = B'\,P'$.

$$(PB')^2 = PC^2 + (CB')^2 = (BB')^2 + PB^2$$

i.e., $c^2{t_1}^2 = v^2{t_1}^2 + d^2$.

$$\therefore \qquad t_1 = \frac{d}{\sqrt{c^2 - v^2}}$$

$\therefore$ Total time taken by the ray to travel the whole path $PB'\,P'$

$$\mathbf{t' = 2t_1 = \frac{2d}{\sqrt{c^2-v^2}} = \frac{2d}{c\sqrt{1-(v^2/c^2)}} = \frac{2d}{c}\left(1+\frac{v^2}{2c^2}\right)} \qquad \text{...(2)}$$

Clearly, $t' < t$. The time difference

$$\Delta t \;=\; t - t' = \frac{2d}{c}\left(1+\frac{v^2}{c^2}\right) - \frac{2d}{c}\left(1+\frac{v^2}{2c^2}\right) = \frac{2d}{c}\times\frac{v^2}{2c^2} = \frac{dv^2}{c^3}.$$

The distance travelled by light in time $\Delta t = c \times \Delta t = dv^2/c^2$.

This is the path difference between the two parts of the incident beam. If the apparatus is turned through 90°, the path difference between the two beams becomes $2dv^2/c^2$. Michelson and Morley expected a fringe shift of about 0.4 in their apparatus when it was rotated through 90° and they believed that they could detect a shift as small as 0.01 of a fringe. But, in the experiment no displacement of the fringes was observed. They repeated the experiment at different points on the earth's surface and at different seasons of the year without detecting any measurable shift in fringes. *This negative result suggests that it is impossible to measure the speed of the earth relative to the ether.* Therefore, the effects of ether are undetectable. Thus, all attempts to make ether as a fixed frame of reference failed.

Explanation of the negative result. The negative result of the Michelson-Morley experiment can be explained by the following three explanations.

(1) The earth dragged along with it the ether in its immediate neighbourhood. Thus, there was no relative motion between the earth and ether. This is the explanation proposed by Michelson himself.

(2) Lorentz and Fitzgerald put forth the suggestion that there was contraction of bodies along the direction of their motion through the ether. Let L_0 be the length of the body when at rest. If it is moving with a speed v parallel to its length, the new length L is $L_0\sqrt{1-(v^2/c^2)}$. In the experiment discussed above, the distance PB will remain unchanged. Distance PA will get shortened to $d\sqrt{1-(v^2/c^2)}$. If d were replaced by $d\sqrt{1-(v^2/c^2)}$ in equation (1), t and t' will be the same and there will be no time difference expected. This contraction hypothesis easily explains why the Michelson-Morley experiment gave a negative result.

(3) The proper explanation for the negative result of the Michelson-Morley experiment was given by Einstein. He concluded that the velocity of light in space is a universal constant. This statement is called the *principle of constancy of the speed of light.* The speed of light is c rather than $|(\mathbf{c} + \mathbf{v})|$ in any frame.

1.7 Special Theory of Relativity

Einstein propounded the special theory of relativity in 1905. The special theory deals with the problems in which one frame of reference moves with a *constant linear velocity* relative to another frame of reference.

Postulates of special theory of relativity

(1) *The laws of Physics are the same in all inertial frames of reference.*

(2) *The velocity of light in free space is constant. It is independent of the relative motion of the source and the observer.*

Explanation :

(1) The first postulate expresses the fact that, since it is impossible to perform an experiment which measures motion relative to a stationary ether, no unique stationary frame of reference can be discovered. Since there is no favoured 'rest' frame of reference, all systems moving with constant

velocity must be on equal footing. We cannot discuss absolute motion. We can discuss only relative motion.

(2) We know that the velocity of light is not constant under Galilean transformations. But according to the second postulate, the velocity of light is the same in all inertial frames. Thus the second postulate is very important and only this postulate is responsible to differentiate the classical theory and Einstein's theory of relativity.

1.8 The Lorentz Transformation Equations

We have to introduce new transformation equations which are consistent with the new concept of the invariance of light velocity in free space. The new transformation equations were discovered by Lorentz, and are known as "Lorentz transformations".

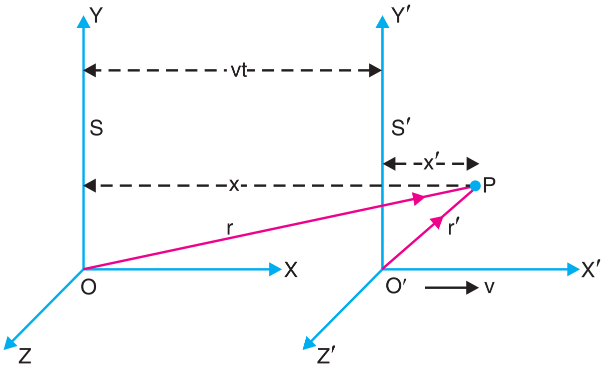


Fig. 1.4

Derivation. Consider two observers O and O' in two systems S and S'. System S' is moving with a constant velocity v relative to system S along the positive X-axis. Suppose we make measurements of time from the instant when the origins of S and S' just coincide, *i.e.*, $t = 0$ when O and O' coincide. Suppose a light pulse is emitted when O and O' coincide. The light pulse produced at $t = 0$ will spread out as a growing sphere. The radius of the wave-front produced in this way will grow with speed c. After a time t, the observer O will note that the light has reached a point P (x, y, z) as shown in Fig. 1.4. For him, the distance of the point P is given by $r = ct$. From figure, $r^2 = x^2 + y^2 + z^2$.

Hence,
$$x^2 + y^2 + z^2 = c^2 t^2. \qquad ...(1)$$

Similarly, the observer O' will note that the light has reached the same point P in a time t' with the same velocity c. So $r' = ct'$.

$$\therefore \qquad x'^2 + y'^2 + z'^2 = c^2 t'^2 \qquad ...(2)$$

Now, equations (1) and (2) must be equal since both the observers are at the centre of the same expanding wavefront.

$$\therefore \qquad x^2 + y^2 + z^2 - c^2 t^2 = x'^2 + y'^2 + z'^2 - c^2 t'^2 \qquad ...(3)$$

Since there is no motion in the Y and Z directions, $y' = y$ and $z' = z$.

$\therefore$ Equation (3) becomes,

$$x^2 - c^2 t^2 = x'^2 - c^2 t'^2 \qquad ...(4)$$

The transformation equation relating to x and x' can be written as,

$$x' = k\,(x - vt) \qquad ...(5)$$

where k is a constant.

The reason for trying the above relation is that, the transformation must reduce to Galilean transformation for low speed ($v << c$).

Similarly, let us assume that

$$t' = a\,(t - bx) \qquad ...(6)$$

where a and b are constants.

Substituting these values for x' and t' in equation (4), we have,

$$x^2 - c^2 t^2 = k^2 (x - vt)^2 - c^2 a^2 (t - bx)^2$$

i.e.,
$$x^2 - c^2 t^2 = (k^2 - c^2 a^2 b^2)\, x^2 - 2(k^2 v - c^2 a^2 b)\, xt - \left(a^2 - \frac{k^2 v^2}{c^2}\right) c^2 t^2 \qquad ...(7)$$

Equating the coefficients of corresponding terms in equation (7),

$$k^2 - c^2 a^2 b^2 = 1 \qquad ...(8)$$

$$k^2 v - c^2 a^2 b = 0 \qquad ...(9)$$

$$a^2 - \frac{k^2 v^2}{c^2} = 1 \qquad ...(10)$$

Solving the above equations for k, a and b, we get,

$$k = a = \frac{1}{\sqrt{1 - (v^2/c^2)}} \qquad ...(11)$$

and
$$b = v/c^2 \qquad ...(12)$$

Substituting these values of k, a and b in (5) and (6) we have,

$$x' = \frac{x - vt}{\sqrt{1 - (v^2/c^2)}} \text{ and } t' = \frac{t - \left(vx/c^2\right)}{\sqrt{1 - \left(v^2/c^2\right)}}$$

Therefore, the Lorentz transformation equations are

$$x' = \frac{x - vt}{\sqrt{1 - (v^2/c^2)}};\ y' = y;\ z' = z \text{ and } t' = \frac{t - \left(vx/c^2\right)}{\sqrt{1 - \left(v^2/c^2\right)}} \qquad ...(13)$$

The inverse Lorentz transformation equations are obtained by interchanging the coordinates and replacing v by $-v$ in the above.

$$x = \frac{x' + vt'}{\sqrt{1 - (v^2/c^2)}},\ y = y';\ z = z' \text{ and } t = \frac{t' + \left(vx'/c^2\right)}{\sqrt{1 - \left(v^2/c^2\right)}} \qquad ...(14)$$

These equations convert measurements made in frame S' into those in frame S.

EXAMPLE 1. *Show that for values of $v << c$, Lorentz transformation reduces to the Galilean transformation.*

SOL. When $v << c$, $\frac{v}{c} \longrightarrow 0$; $\therefore \frac{1}{\sqrt{1 - (v^2/c^2)}} \approx 1.$

We have from (13), $x' = x - vt$; $y' = y$; $z' = z$ and $t' = t$

which are Galilean transformations.

EXAMPLE 2. *Show that if (x_1, y_1, z_1, t_1) and (x_2, y_2, z_2, t_2) are the coordinates of one event in S_1 and the corresponding event in S_2 respectively, then the expression*

$$ds_1^2 = dx_1^2 + dy_1^2 + dz_1^2 - c^2\, dt_1^2$$

is invariant under a Lorentz transformation of coordinates.

SOL. The inverse Lorentz transformation equations are:

$$x_1 = \frac{x_2 + vt_2}{\sqrt{1 - \beta^2}};\ y_1 = y_2;\ z_1 = z_2 \text{ and } t_1 = \frac{t_2 + vx_2/c^2}{\sqrt{1 - \beta^2}} \quad (\text{where } v/c = \beta).$$

Differentiating,

$$dx_1 = \frac{dx_2 + v\,dt_2}{\sqrt{1-\beta^2}};\ dy_1 = dy_2;\ dz_1 = dz_2 \ \text{ and } \ dt_1 = \frac{dt_2 + (\beta/c)\,dx_2}{\sqrt{1-\beta^2}}$$

$$ds_1{}^2 = \left(\frac{dx_2 + v\,dt_2}{\sqrt{1-\beta^2}}\right)^2 + dy_2{}^2 + dz_2{}^2 - c^2\left(\frac{dt_2 + (\beta/c)\,dx_2}{\sqrt{1-\beta^2}}\right)^2$$

which simplifies to

$$ds_1{}^2 = dx_2{}^2 + dy_2{}^2 + dz_2{}^2 - c^2\,dt_2{}^2 = ds_2{}^2$$

1.9 Length Contraction

Consider two coordinate systems S and S' with their X-axes coinciding at time $t = 0$. S' is moving with a uniform relative speed v with respect to S in the positive X-direction. Imagine a rod (AB), at rest relative to S' (Fig. 1.5).

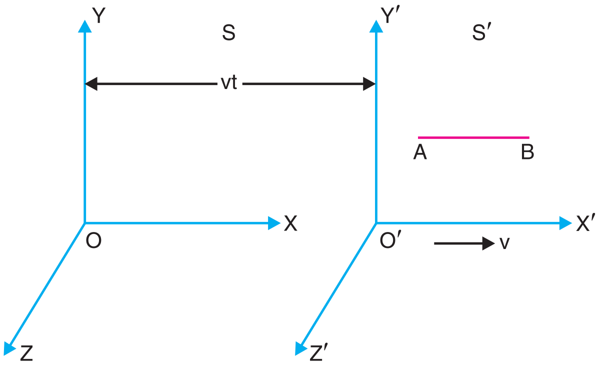


Fig. 1.5

Let x'_1 and x'_2 be the coordinates of the ends of the rod at any instant of time in S'. Then,

$$l_0 = x_2' - x_1' \qquad \text{...(1)}$$

since the rod is at rest in frame S'.

Similarly, let x_1 and x_2 be the coordinates of the ends of the rod at the same instant of time in S.

Then $$l = x_2 - x_1 \qquad \text{...(2)}$$

l the length of the rod, measured relative to S.

According to Lorentz transformations,

$$x_2' = \frac{x_2 - vt}{\sqrt{1-(v^2/c^2)}} \qquad \text{...(3)}$$

and $$x_1' = \frac{x_1 - vt}{\sqrt{1-(v^2/c^2)}} \qquad \text{...(4)}$$

Illustration of Einstein's Theory of Length Contraction.

Subtracting equation (4) from (3)

$$x_2' - x_1' = \frac{x_2 - x_1}{\sqrt{1-(v^2/c^2)}} \quad \text{or} \quad l_0 = \frac{l}{\sqrt{1-(v^2/c^2)}}$$

$$\therefore \qquad l = l_0 \sqrt{1-(v^2/c^2)} \qquad \text{...(5)}$$

From equation (5) we see that $l < l_0$. Therefore, to the observer in S it would appear that the length of the rod (in S') has contracted by the factor $\sqrt{1-(v^2/c^2)}$.

For example, a body which appears to be spherical to an observer at rest relative to it, will appear to be an oblate spheroid to a moving observer. Similarly, a square and a circle in one appear to the observer in the other to be a rectangle and an ellipse respectively (Fig. 1.6).

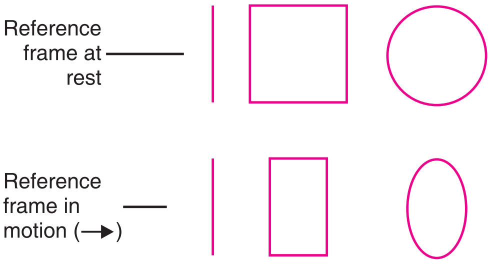


Fig.1.6

Notes

(1) The *proper length* of an object is the length determined by an observer at rest with respect to the object. In the above case, l_0 is the proper length.

(2) The shortening or contraction in the length of an object along its direction of motion is known as the Lorentz-Fitzgerald contraction. There is no contraction in a direction perpendicular to the direction of motion.

(3) The contraction becomes appreciable only when $\mathbf{v} \approx \mathbf{c}$.

(4) The contraction is reciprocal, *i.e.*, if two identical rods are at rest—one in S and the other in S', each of the observers finds that the other is shorter than the rod of his own system.

EXAMPLE 1. *A rod* 1 *metre long is moving along its length with a velocity* 0.6*c. Calculate its length as it appears to an observer* (*a*) *on the earth* (*b*) *moving with the rod itself.*

SOL. Here, 1 metre is the proper length (l_0) of the rod in its own moving frame of reference. (*a*) Let l be the length of the rod as it appears to an observer in the stationary reference frame of the earth.

Here, $\quad \boldsymbol{l_0 = 1\ \mathrm{m};\ v = 0.6\ c;\ l = ?}$

$$l = l_0\sqrt{1-\frac{v^2}{c^2}} = 1\sqrt{1-\frac{(0.6c)^2}{c^2}} = 1\sqrt{1-0.36} = 0.8\ \text{m}$$

Hence, the observer on the earth will estimate the length of the rod to be 0.8 metre.

(*b*) For an observer moving with the rod itself, the length of the rod is 1 metre.

EXAMPLE 2. *How fast would a rocket have to go relative to an observer for its length to be contracted to 99% of its length at rest?*

SOL. Here, $\quad l = 0.99 l_0; \quad v = ?$

We have, $\quad l = l_0\sqrt{1-(v^2/c^2)}$

$$l^2 = l_0^2\left(1-\frac{v^2}{c^2}\right) \quad \text{or} \quad \left(1-\frac{v^2}{c^2}\right) = \frac{l^2}{l_0^2} \quad \text{or} \quad \frac{v^2}{c^2} = 1-\frac{l^2}{l_0^2}.$$

$$\therefore \qquad v^2 = c^2\left(1-\frac{l^2}{l_0^2}\right) = c^2\,(1-0.99^2)$$

$$\therefore \qquad v = 0.1416\,c = 0.1416 \times (3 \times 10^8) = 4.245 \times 10^7\ \text{ms}^{-1}$$

EXAMPLE 3. *A rigid bar of length $L_2 = 1.5$ m is at rest relative to system S' (Fig. 1.7). If the bar makes an angle $\theta_2 = 45°$ with respect to the x_2 axis, what is the length L_1 and orientation of the bar θ_1 relative to S when $v = 0.98c$?*

SOL. When the length of the bar is resolved into the two components parallel to the x_2 and y_2 axes, respectively, the corresponding lengths measured by S' will be

$$L_{2x} = L_2 \cos\theta_2$$

$$L_{2y} = L_2 \sin\theta_2$$

The vertical component is perpendicular to v and will not experience any contraction when viewed from S. Hence,

$$L_{1y} = L_{2y} = L_2 \sin\theta_2.$$

The horizontal component is parallel to v and will be contracted according to equation (5) to give

$$L_{1x} = L_{2x}\sqrt{1-(v^2/c^2)} = L_2\sqrt{1-(v^2/c^2)}\cos\theta_2.$$

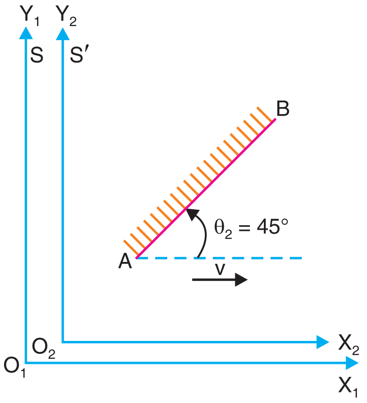


Fig. 1.7

The length of the bar as measured by O_1 will be

$$L_1 = \sqrt{L_{1x}{}^2 + L_{1y}{}^2} = \sqrt{L_2^2(1-v^2/c^2)\cos^2\theta_2 + L_2^2\sin^2\theta_2}$$

$$= L_2\sqrt{1-(v^2/c^2)\cos^2\theta_2} = 1.5\sqrt{1-\left\{(0.98c)^2/c^2\right\}\cos^2 45°} = 1.08m\cdot$$

The orientation relative to S will be given by

$$\tan\theta_1 = \frac{L_{1y}}{L_{1x}} = \frac{L_2\sin\theta_2}{L_2\sqrt{1-(v^2/c^2)}\cos\theta_2}$$

$$= \frac{\tan\theta_2}{\sqrt{1-(v^2/c^2)}} = \frac{\tan 45°}{\sqrt{1-(0.98\,c)^2/c^2}}$$

$$\therefore \qquad \theta_1 = 78.7°$$

1.10 Time Dilation

Imagine a gun placed at the position (x', y', z') in S'. Suppose it fires two shots at times t_1' and t_2' measured with respect to S'. In S' the clock is at rest relative to the observer. The time interval measured by a clock at rest relative to the observer is called the *proper time interval*. Hence, $t_0 = t_2' - t_1'$ is the time interval between the two shots for the observer in S'. Since the gun is fixed in S', it has a velocity v with respect to S in the direction of the positive X-axis. Let $t = t_2 - t_1$ represent the time interval between the two shots as measured by an observer in S.

From inverse Lorentz transformations we have

$$t_1 = \frac{t_1' + vx'/c^2}{\sqrt{1-(v^2/c^2)}} \quad \text{and} \quad t_2 = \frac{t_2' + vx'/c^2}{\sqrt{1-(v^2/c^2)}}$$

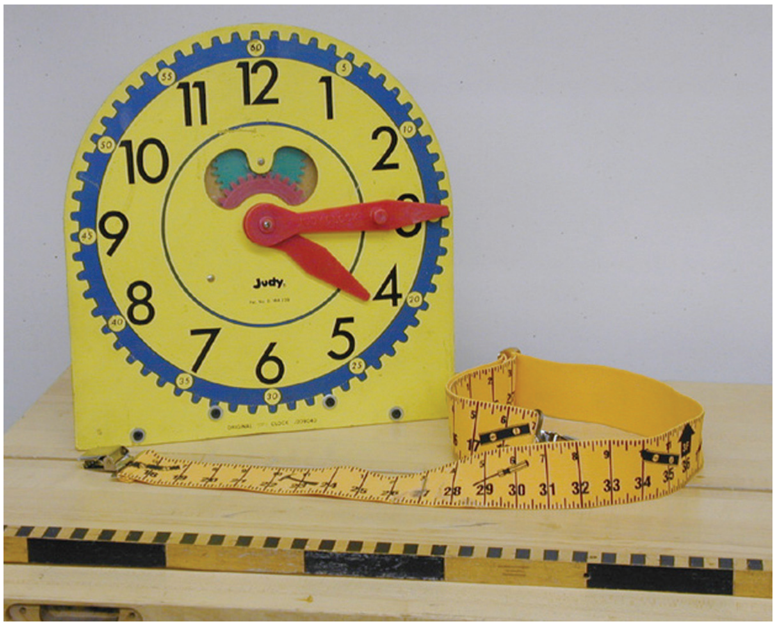
$$\therefore \qquad t_2 - t_1 = \frac{t_2' - t_1'}{\sqrt{1-(v^2/c^2)}} \quad \text{or} \quad t = \frac{t_0}{\sqrt{1-(v^2/c^2)}}$$

$$\text{or} \qquad t > t_0.$$

Thus, the time interval, between two events occurring at a given point in the moving frame S' appears to be longer to the observer in the stationary frame S; *i.e.*, a stationary clock measures a longer time interval between events occurring in a moving frame of reference than does a clock in the moving frame. This effect is called *time dilation.*

The Twin Paradox. Consider two exactly identical twin brothers. Let one of the twins go to a long space journey at a high speed in a rocket and the other stay behind on the earth. The clock in the moving rocket will appear to go slower than the clock on the surface of the earth (in accordance with $t = \dfrac{t_0}{\sqrt{1-(v^2/c^2)}}$). Therefore, when he returns back to the earth, he will find himself younger than the twin who stayed behind on the earth!

Illustration of time dilation and length contraction — Meson decay. A striking illustration of both the time dilation and length contraction occurs in the case of unstable particles called μ mesons. A μ meson decays into an electron an average of 2×10^{-6} sec. after it comes into being, *i.e.*, the meson's mean lifetime is 2×10^{-6} sec. Now μ mesons are created high in the atmosphere by fast cosmic ray particles arriving at the earth from space, and reach sea level in profusion. Such mesons have a speed of 2.994×10^8 ms^{-1} which is $0.998c$. But in $t_0 = 2 \times 10^{-6}$ s = the meson's mean lifetime, they can travel a distance of only $y = vt_0 = (2.994 \times 10^8)(2 \times 10^{-6}) = 600$ m. But, the μ mesons are actually created at altitudes more than 10 times greater than this.



Time Dilation/Length Contraction.

We can resolve the meson paradox by using the results of the special theory of relativity. Let us examine the problem from the frame of reference of the meson, in which its lifetime is 2×10^{-6} s. While the meson lifetime is unaffected by the motion, its distance to the ground appears shortened by the factor $y/y_0 = \sqrt{1-v^2/c^2}$. That is, while we, on the ground, measure the altitude at which the meson is produced as y_0, the meson "sees" it as y. Let y = the maximum distance the meson can travel in its own frame of reference at the speed $0.998c$ before decaying. Then corresponding distance y_0 in our reference frame is

$$y_0 = \frac{y}{\sqrt{1-\frac{v^2}{c^2}}} = \frac{600}{\sqrt{1-\frac{(0.998c)^2}{c^2}}} = \frac{600}{\sqrt{1-0.996}} = \frac{600}{0.063} = 9500\,\text{m}.$$

Hence, despite their brief lifespans, it is possible for the mesons to reach the ground from the considerable altitudes at which they are actually formed.

Now let us examine the problem from the frame of reference of an observer on the ground. From the ground the altitude at which the meson is produced is y_0. But the lifetime of the meson in *our reference frame* has been extended, owing to the relative motion, to the value t.

$$t = \frac{t_0}{\sqrt{1-\frac{v^2}{c^2}}} = \frac{2\times10^{-6}}{\sqrt{1-\frac{(0.998c)^2}{c^2}}} = \frac{2\times10^{-6}}{0.063} = 31.7 \times 10^{-6}\text{ s}$$

$\therefore$ Distance travelled by the meson in 31.7×10^{-6} sec.

$$y_0 = vt = (2.994 \times 10^8)(31.7 \times 10^{-6}) = 9500\text{ m}.$$

This is the same as the distance obtained before.

Thus the two points of view give identical results.

EXAMPLE 1. *A clock in a space ship emits signals at intervals of* 1 *second as observed by an astronaut in the space ship. If the space ship travels with a speed of* 3×10^7 *ms*$^{-1}$, *what is the interval between successive signals as seen by an observer at the control centre on the ground*?

SOL. Here, $t_0 = 1\text{s};\ v = 3 \times 10^7\text{ ms}^{-1};$ and $c = 3 \times 10^8\text{ ms}^{-1};\ t = ?$

$$\therefore \quad t = \frac{t_0}{\sqrt{1-\frac{v^2}{c^2}}} = \frac{1}{\sqrt{1-\frac{(3\times10^7)^2}{(3\times10^8)^2}}} = 1.005\text{ s}.$$

EXAMPLE 2. *A particle with a proper lifetime of* 1μs *moves through the laboratory at* 2.7×10^8 ms^{-1}. (*a*) *What is its lifetime, as measured by observers in the laboratory*? (*b*) *What will be the distance traversed by it before disintegrating*?

SOL. Here, $t_0 = 1\mu s = 10^{-6}\text{s};\ v = 2.7 \times 10^8\text{ ms}^{-1};\ t = ?$

(*a*)
$$t = \frac{t_0}{\sqrt{1-\frac{v^2}{c^2}}} = \frac{10^{-6}}{\sqrt{1-\frac{(2.7\times10^8)^2}{(3\times10^8)^2}}} = 2.3 \times 10^{-6}\text{ s}.$$

(*b*) The average distance moved by the particle before disintegration $= (2.7 \times 10^8) \times (2.3 \times 10^{-6})$
$= 620$ m.

1.11 Relativity of Simultaneity

Consider two events—the explosion of a pair of time bombs—that occur at the same time to an observer O in a reference frame S. Let the two events occur at different locations x_1 and x_2. Consider another observer O' in S' moving with a uniform relative speed v with respect to S in the positive X-direction.

To O', the explosion at x_1 and t_0 occurs at the time

$$t_1' = \frac{t_0 - (v/c^2)x_1}{\sqrt{1-(v^2/c^2)}}$$

and the explosion at x_2 and t_0 occurs at the time

$$t'_2 = \frac{t_0 - (v/c^2)x_2}{\sqrt{1-(v^2/c^2)}}$$

$\therefore$ The time interval between the two events as observed by the observer O'

$$= t' = t'_2 - t'_1 = \frac{(v/c^2)(x_1 - x_2)}{\sqrt{(1-v^2/c^2)}}$$

This is not zero. This indicates that two events at x_1 and x_2, which are simultaneous to the observer in S, do not appear so to the observer in S'. Therefore, the concept of simultaneity has only a relative and not an absolute meaning.

1.12 Addition of Velocities

Suppose the system S' moves with a uniform velocity v relative to the system S. Suppose a particle is moving in the common direction of the x and x' axes. Let its velocity, as measured by an observer in the system S, be u and as measured by an observer in S' be u'.

Then we have, $u = \dfrac{dx}{dt}$ and $u' = \dfrac{dx'}{dt}$.

Using inverse Lorentz transformations, we have, $x = \dfrac{x'+vt'}{\sqrt{1-(v^2/c^2)}}$ and $t = \dfrac{t'+vx'/c^2}{\sqrt{1-(v^2/c^2)}}$.

Differentiating, $$dx = \frac{dx'+v\,dt'}{\sqrt{1-(v^2/c^2)}} \quad \text{and} \quad dt = \frac{dt'+v\,dx'/c^2}{\sqrt{1-(v^2/c^2)}}$$

Thus $$\frac{dx}{dt} = \frac{dx'+v\,dt'}{dt'+\dfrac{v}{c^2}dx'} = \frac{\left(\dfrac{dx'}{dt'}\right)+v}{1+\dfrac{v}{c^2}\left(\dfrac{dx'}{dt'}\right)}$$

$\therefore$ $$u = \frac{u'+v}{1+u'v/c^2}.$$

If we put $u' = c$ and $v = c$ *i.e.*, photon is moving with a velocity c in frame S' and S' is moving with velocity c relative to S, then

$$u = \frac{c+c}{1+(c^2/c^2)} = c.$$

Thus, the addition of any velocity to the velocity of light c merely reproduces the velocity of light. Hence, the velocity of light is the maximum attainable velocity.

Notes

(1) This law of addition of velocities applies only when the two velocities are in the same direction.

(2) If $v << c$, we get the classical equation.

(3) We can express the velocity u' in terms of u and v.

$$u' = \frac{u-v}{1-\dfrac{uv}{c^2}}$$

Example 1. *An experimenter observes a radioactive atom moving with a velocity of* 0.25*c*. *The atom then emits a* β *particle which has a velocity of* 0.9*c relative to the atom in the direction of its motion. What is the velocity of the* β *particle, as observed by the experimenter*?

Sol. Here, $v = 0.25c;\ u' = 0.9c;\ u = ?$

$$u = \frac{u'+v}{1+\frac{v}{c^2}u'} = \frac{0.9c+0.25c}{1+\frac{0.25c}{c^2}\times 0.9c} = 0.94c.$$

Example 2. *An electron is moving with a speed of* 0.85*c in a direction opposite to that of a moving photon. Calculate the relative velocity of the photon with respect to the electron.*

Sol. Let the photon be moving along the positive direction of *X*-axis and the electron along the negative direction of *X*-axis. Then, the speed of electron = – 0.85*c*; the speed of photon = *c*. Consider that the electron is at rest in system *S*. Then, we may assume that the system *S'* (laboratory) is moving with velocity 0.85*c* relative to system *S* (electron). *i.e.*, $v = 0.85c;\ u' = c;\ u = ?$

$$u = \frac{u'+v}{1+\frac{u'v}{c^2}} = \frac{c+0.85c}{1+\frac{(c)\times(0.85c)}{c^2}} = c.$$

1.13 Variation of Mass with Velocity

Derivation. Consider two systems *S* and *S*′. *S*′ is moving with a constant velocity *v* relative to the system *S*, in the positive *X*-direction (Fig. 1.8). Suppose that in the system *S*′, two exactly similar elastic balls *A* and *B* approach each other at equal speeds (*i.e.*, *u* and –*u*). Let the mass of each ball be *m* in *S*′. They collide with each other and after collision coalesce into one body. According to the law of conservation of momentum,

Momentum of ball *A* + momentum of ball *B* = momentum of coalesced mass.

Or $mu + (-mu)$ = momentum of coalesced mass = 0.

Thus the coalesced mass must be at rest in *S*′ system.

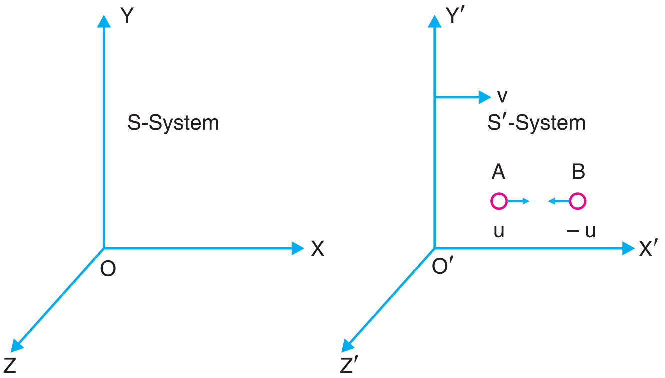


Fig. 1.8

Let us now consider the collision with reference to the system *S*.

Let u_1 and u_2 be the velocities of the balls relative to *S*. Then,

$$u_1 = \frac{u+v}{1+uv/c^2} \qquad \text{...(1)}$$

and

$$u_2 = \frac{-u+v}{1-uv/c^2} \qquad \text{...(2)}$$

After collision, velocity of the coalesced mass is v relative to the system S.

Let mass of the ball A travelling with velocity u_1 be m_1 and that of B with velocity u_2 be m_2 in the system S. Total momentum of the balls is conserved. Therefore,

$$m_1 u_1 + m_2 u_2 = (m_1 + m_2)\, v \qquad \text{...(3)}$$

Substituting for u_1 and u_2 from equations (1) and (2), we have,

$$m_1\left[\frac{u+v}{1+uv/c^2}\right] + m_2\left[\frac{-u+v}{1-uv/c^2}\right] = (m_1 + m_2)v$$

or
$$m_1\left[\frac{u+v}{1+\dfrac{uv}{c^2}} - v\right] = m_2\left[v - \frac{-u+v}{1-\dfrac{uv}{c^2}}\right]$$

or
$$m_1\left[\frac{u+v-v-uv^2/c^2}{1+uv/c^2}\right] = m_2\left[\frac{v-uv^2/c^2+u-v}{1-uv/c^2}\right]$$

or
$$m_1\left[\frac{u\left(1-\dfrac{v^2}{c^2}\right)}{1+\dfrac{uv}{c^2}}\right] = m_2\left[\frac{u\left(1-\dfrac{v^2}{c^2}\right)}{1-\dfrac{uv}{c^2}}\right]$$

or
$$\frac{m_1}{m_2} = \frac{1+uv/c^2}{1-uv/c^2} \qquad \text{...(4)}$$

Also,
$$1-\frac{u_1^2}{c^2} = 1 - \frac{\left\{\dfrac{u+v}{c}\right\}^2}{\left(1+\dfrac{uv}{c^2}\right)^2}$$

$$= \frac{1+\dfrac{u^2v^2}{c^4}+\dfrac{2uv}{c^2}-\dfrac{u^2}{c^2}-\dfrac{v^2}{c^2}-\dfrac{2uv}{c^2}}{\left(1+\dfrac{uv}{c^2}\right)^2}$$

$$= \frac{\left(1-\dfrac{u^2}{c^2}\right)-\dfrac{v^2}{c^2}\left(1-\dfrac{u^2}{c^2}\right)}{\left(1+\dfrac{uv}{c^2}\right)^2}$$

$$1-\frac{u_1^2}{c^2} = \frac{\left(1-\dfrac{u^2}{c^2}\right)\left(1-\dfrac{v^2}{c^2}\right)}{\left(1+\dfrac{uv}{c^2}\right)^2} \qquad \text{...(5)}$$

Similarly,
$$1-\frac{u_2^2}{c^2} = \frac{\left(1-\dfrac{u^2}{c^2}\right)\left(1-\dfrac{v^2}{c^2}\right)}{\left(1-\dfrac{uv}{c^2}\right)^2} \qquad \text{...(6)}$$

Dividing equation (6) by equation (5),

$$\frac{1-\dfrac{u_2^2}{c^2}}{1-\dfrac{u_1^2}{c^2}} = \frac{\left(1+\dfrac{uv}{c^2}\right)^2}{\left(1-\dfrac{uv}{c^2}\right)^2}$$

or

$$\frac{\sqrt{1-\dfrac{u_2^2}{c^2}}}{\sqrt{1-\dfrac{u_1^2}{c^2}}} = \frac{1+\dfrac{uv}{c^2}}{1-\dfrac{uv}{c^2}} \qquad \text{...(7)}$$

From equations (7) and (4), $\dfrac{m_1}{m_2} = \dfrac{\sqrt{1-\dfrac{u_2^2}{c^2}}}{\sqrt{1-\dfrac{u_1^2}{c^2}}}$

or

$$m_1\sqrt{1-\frac{u_1^2}{c^2}} = m_2\sqrt{1-\frac{u_2^2}{c^2}} \qquad \text{...(8)}$$

Since the L.H.S. and R.H.S. of equation (8) are independent of one another, the above result can be true only if each is a constant. Therefore,

$$m_1\sqrt{1-\frac{u_1^2}{c^2}} = m_2\sqrt{1-\frac{u_2^2}{c^2}} = m_0.$$

The constant denoted by m_0 is called the *rest mass* of the body and corresponds to *zero velocity.*

Thus,

$$m_1 = \frac{m_0}{\sqrt{1-u_1^2/c^2}}$$

In general, if m denotes the mass of a body when it is moving with a velocity v, then,

$$m = \frac{m_0}{\sqrt{1-v^2/c^2}} \qquad \text{...(9)}$$

This is the relativistic formula for the variation of mass with velocity.

If we put $v \to c$ in equation (9), we have $m \to \infty$ *i.e.*, an object travelling at the speed of light would have infinite mass. Thus, Eqn. (9) shows that no material body can have a velocity equal to, or greater than the velocity of light.

Notes

(1) The first verification of the increase in mass with velocity came from the work of Kaufmann in 1906 and of Bucherer in 1909. While studying the β-rays emanating from radioactive materials, they found that their velocities were comparable to the velocity of light and also that their masses were found to be related to their velocities.

(2) The increase of mass with velocity has now been tested in "*particle accelerators*". Accelerators are used to accelerate the various charged particles of matter to very high velocities. It has been found that the electrons and protons accelerated in these machines to velocities close to the velocity of light acquire increased masses, exactly as predicted.

One Million Volt
Particle Accelerator
1937

Particle Accelerator.

Example. *At what speed is a particle moving if the mass is equal to three times its rest mass?*

Sol. We have, $m = \frac{m_0}{\sqrt{1-v^2/c^2}}$. Here, $m = 3m_0$; $v = ?$

$$\therefore \quad 3m_0 = \frac{}{\sqrt{1 \quad v^2/c^2}} \text{ or } \sqrt{1-\frac{v^2}{c^2}} = \frac{1}{3} \text{ or } 1-\frac{v^2}{c^2} = \frac{1}{9}$$

Or $$\frac{v^2}{c^2} = \frac{8}{9} \text{ or } v^2 = \frac{8}{9}c^2 \text{ or } v = \sqrt{\frac{8}{9}}c = 0.94c.$$

1.14 Mass Energy Equivalence

Derivation. Force is defined as rate of change of momentum *i.e.*,

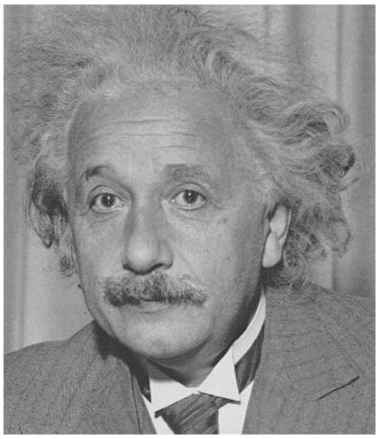
$$F = \frac{d}{dt}(mv) \qquad ...(1)$$

According to the theory of relativity, both mass and velocity are variable. Therefore,

$$F = \frac{d}{dt}(mv) = m\frac{dv}{dt}+v\frac{dm}{dt} \qquad ...(2)$$

Let the force F displace the body through a distance dx.

Then, the increase in the kinetic energy (dE_k) of the body is equal to the work done ($F\,dx$).



Albert Einstein (1879-1955)

Hence, $$dE_k = F\,dx = m\frac{dv}{dt}dx + v\frac{dm}{dt}dx$$

or $$dE_k = mv\,dv + v^2\,dm \qquad ...(3)$$

According to the law of variation of mass with velocity

$$m = \frac{m_0}{\sqrt{(1-v^2/c^2)}} \qquad ...(4)$$

Squaring both sides, $$m^2 = \frac{m_0{}^2}{1-(v^2/c^2)}$$

or $$m^2c^2 = m_0^2c^2 + m^2v^2$$

Differentiating, $$c^2\,2m\,dm = m^2\,2v\,dv + v^2\,2m\,dm$$

or $$c^2\,dm = mv\,dv + v^2\,dm \qquad ...(5)$$

From equations (3) and (5), $$dE_k = c^2\,dm \qquad ...(6)$$

Thus, a change in K.E. dE_k is directly proportional to a change in mass dm.

When a body is at rest, its velocity is zero, (K.E. = 0) and $m = m_0$. When its velocity is v, its mass becomes m. Therefore, integrating equation (6),

$$E_k = \int_o^{E_k} dE_k = c^2\int_{m_o}^{m} dm = c^2(m-m_0)$$

$$\therefore \quad E_k = mc^2 - m_oc^2 \qquad ...(7)$$

This is the relativistic formula for K.E.

When the body is at rest, the internal energy stored in the body is m_0c^2. m_0c^2 is called the rest mass energy. The total energy (E) of the body is the sum of K.E. (E_k) and rest mass energy (m_0c^2).

$$\therefore \qquad E = E_k + m_0c^2 = (mc^2 - m_0c^2) + m_0c^2 = mc^2.$$

$$\therefore \qquad E = mc^2$$

This is Einstein's mass-energy relation.

Notes

(1) This relation states a universal equivalence between mass and energy. It means that mass may appear as energy and energy as mass.

Example. Consider the phenomenon of pair-annihilation or pair production. In this phenomenon, an electron and a positron can combine and literally disappear. In their place we find high energy radiation called γ-radiation, whose radiant energy is exactly equal to the rest mass plus kinetic energies of the disappearing particles. The process is reversible, so that a materialization of mass from radiant energy can occur when a high enough energy γ-ray, under proper conditions, disappears. In its place appears a positron-electron pair whose total energy (rest mass + K.E.) is equal to the radiant energy lost.

(2) The relationship ($E = mc^2$) between energy and mass forms the basis of understanding nuclear reactions such as *fission* and *fusion*. These reactions take place in nuclear bombs and reactors. When a uranium nucleus is split up, the decrease in its total rest mass appears in the form of an equivalent amount of K.E. of its fragments.

(3) The formula for K.E. reduces to the classical formula for $v << c$.

$$E_k = mc^2 - m_0c^2 = (m - m_0)\,c^2 = m_0\,c^2\,[(1 - v^2/c^2)^{-½} - 1]$$

Now, since $v<<c, \left(1-\dfrac{v^2}{c^2}\right)^{-\frac{1}{2}} = 1+\dfrac{v^2}{2c^2} + \ldots$ (neglecting higher order terms).

$$\therefore \qquad E_k = m_0c^2 \times \frac{v^2}{2c^2} = \frac{1}{2}m_0v^2.$$

Unified mass unit. The basic mass unit used in Atomic Physics is called the unified mass unit (u). It is defined as (1/12)th of the rest mass of carbon—12 atom. Thus, $1u = 1.66 \times 10^{-27}$ kg.

Now, $$E = m_0\,c^2 = (1.66 \times 10^{-27})\,(3 \times 10^8)^2 = 1.49 \times 10^{-10}\text{ J.}$$

$$= \frac{1.49 \times 10^{-10}}{1.6 \times 10^{-19}}\,eV = 931.3\text{ MeV.}$$

Hence the energy equivalent to unified mass unit is 931 MeV.

EXAMPLE 1. *If 4 kg of a substance is fully converted into energy, how much energy is produced?*

SOL. Here, $$m = 4\text{ kg}, c = 3 \times 10^8\text{ ms}^{-1}; E = ?$$

$$E = mc^2 = 4 \times (3 \times 10^8)^2 = 3.6 \times 10^{17}\text{ J.}$$

EXAMPLE 2. *Calculate the rest energy of an electron in joules and in electron volts.*

SOL. Here, m_0 = rest mass of the electron = 9.11×10^{-31} kg;

$$c = 3 \times 10^8\text{ ms}^{-1}$$

$$\therefore \qquad E = m_0c^2 = (9.11 \times 10^{-31})\,(3 \times 10^8)^2 = 8.2 \times 10^{-14}\text{ J.}$$

$$= \frac{8.2 \times 10^{-14}}{1.6 \times 10^{-19}}\,eV \quad (\text{since } 1eV = 1.6 \times 10^{-19}\ J).$$

$$= 0.51\text{ MeV [mega electron volt].}$$

EXAMPLE 3. *Calculate the K.E. of an electron moving with a velocity of* 0.98 *times the velocity of light in the laboratory system.*

SOL. Relativistic formula for K.E. is $T = (m - m_0)\, c^2$.

Here, m_0 = rest mass of electron = 9.11×10^{-31} kg.

$c = 3 \times 10^8 \text{ ms}^{-1}$; $v = 0.98c$.

We can find m using the formula

$$m = \frac{m_0}{\sqrt{(1 - v^2/c^2)}} = \frac{m_0}{\sqrt{1-(0.98c/c)^2}} = 5.02$$

$$\therefore \quad \text{K.E.} = (5.02\, m_0 - m_0)c^2 = 4.02\, m_0c^2$$

$$= 4.02 \times (9.11 \times 10^{-31})\,(3 \times 10^8)^2 = 3.296 \times 10^{-13}\, J.$$

Relationship between the total energy, the rest energy, and the momentum :

The total relativistic energy (rest mass plus kinetic) of a particle is

$$E = mc^2 = m_0c^2 / (1 - v^2/c^2)^{1/2} \qquad \text{...(1)}$$

The momentum of the particle is $p = mv$

or $v = p/m.$ ...(2)

$$\therefore \quad E = m_0 \frac{c^2}{\sqrt{[1-(p^2/m^2c^2)]}} = \frac{m_0c^2}{\sqrt{[1-(p^2c^2/m^2c^4]}} = m_0 \frac{c^2}{\sqrt{[1-(p^2c^2/E^2)]}}$$

$$\text{or} \quad E^2 = m_0^{\,2} \frac{c^4}{(1-p^2c^2/E^2)} \quad \text{or} \quad E^2 - p^2c^2 = m_0^{\,2}c^4$$

$$\therefore \quad E^2 = m_0^{\,2}\, c^4 + p^2c^2 \qquad \text{...(3)}$$

1.15 Minkowski's Four Dimensional Space-Time Continuum

Space is three dimensional in the sense that it is possible to describe the position of a point by three coordinate numbers as in Euclidean geometry. But the truth is that the world in which we live is a four-dimensional space-time continuum. This is revealed by the theory of relativity.

Any event in any frame of reference S can be represented by the four coordinates x, y, z, t. Suppose the same event is observed in another frame S' moving with uniform velocity v relative to S along the X-axis and coinciding with S at $t = t' = 0$. Then its coordinates in S' will be x', y', z', t', satisfying the equation

$$x^2 + y^2 + z^2 - c^2 t^2 = x'^2 + y'^2 + z'^2 - c^2 t'^2 \qquad \text{...(1)}$$

It is possible to characterise the Lorentz transformations more elegantly if we insert $x_1 = x$, $x_2 = y$, $x_3 = z$ and $x_4 = \sqrt{-1}\, ct = ict$ for the system S and similarly for S'. The resulting form of equation (1) is

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = x_1'^2 + x_2'^2 + x_3'^2 + x_4'^2 \qquad \text{...(2)}$$

The imaginary time coordinate x_4 appears here in exactly the same way as the space coordinates x_1, x_2 and x_3. This four dimensional manifold which appears as a linking together of space and time is known as the "*Minkowski world*" or the 'four dimensional space-time continuum'. Here the *four coordinates* are x, y, z and ict, and time is equivalent to an imaginary length. The world of Minkowski involves space and time inextricably bound with each other. A point in such a space-time system is known as the "*world point*", and the motion of a particle in it may be represented by a "*world line*".

If s be the distance between two points with co-ordinates (x_1, x_2, x_3, x_4) and (x_1', x_2', x_3', x_4') in the four dimensional space, then

$$s^2 = (x_1' - x_1)^2 + (x_2' - x_2)^2 + (x_3' - x_3)^2 + (x_4' - x_4)^2$$

This can be shown to remain invariant *i.e.*, unaltered in form in all systems by the use of the transformation equations. For a "*worldline element*" this invariant form of equation is

$$ds^2 = dx_1^2 + dx_2^2 + dx_3^2 + dx_4^2 = dx^2 + dy^2 + dz^2 - c^2\, dt^2$$

This means that the distance between two points, *i.e.*, a length, is not defined by the three dimensions dx, dy, dz alone, but in addition, by a fourth dimension of time dt.

For ds = minimum, we get a straight line between two world points.

This straight world line is called a *geodesic*, for which

$$\delta \int_A^B ds = 0.$$

1.16 The General Theory of Relativity

The general theory of relativity is based on the following two postulates.

Postulate 1. *"The laws of physics may be expressed in equations having the same form in all frames of reference, regardless of their states of motion".*

Thus the general theory covers accelerated as well as uniform motion. Hence it is able to describe gravitational phenomena.

Postulate 2. *There is no way for an observer in a closed laboratory to distinguish between the effects produced by a gravitational field and those produced by an acceleration of the laboratory.*

This postulate is known as the *principle of equivalence.* This postulate follows from the experimental observation that the inertial mass of a body is always exactly equal to its gravitational mass.

Einstein generalised the principle of equivalence and proposed that the physical laws have similar mathematical forms not only in inertial systems but in accelerated systems also. According to him, the idea of action at a distance for exploration of gravitation is false and by selecting a suitable coordinate system, the effect of gravitation can be eliminated. In this system, the inertial mass of a body is equal to its gravitational mass. To explain these ideas, let us consider the weight of a body in a lift. If the lift goes upwards with an acceleration a, the weight of the body is $m(g + a)$. If the lift goes down with acceleration a, the weight of the body is $m(g - a)$. If the lift goes down with acceleration g, the weight of the body will be zero. Hence the gravitational effect on the body depends on the state of the lift.

1.17 Predictions of General Relativity

1. Deflection of light in a gravitational field

Inertial mass $h\nu/c^2$ is associated with a light pulse of energy $h\nu$. *From the principle of equivalence, the gravitational mass of this quantum is* $h\nu/c^2$ so that light ought to be subject to gravitational forces. For example, light rays that pass near the sun ought to be deflected towards it, just as the paths of the planets are deflected by the sun's gravitational field. This deflection for a ray of light is shown in Fig. 1.9. The deflection is known as *'Einstein effect'*. This effect can be detected at the time of total solar eclipse. The experimental value of the observed deflection is almost equal to its theoretical value which is 1.75 seconds.

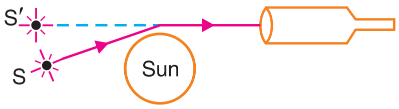


Fig. 1.9

2. Advance of the perihelion of Mercury's orbit

The general theory predicts that the path of a planet about the sun will not be precisely an ellipse. The orbit will slowly rotate in the plane of its motion, so that the axes turn in space in the

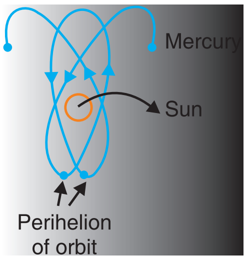


Fig. 1.10

same sense as the planet's rotation. For the planets in our solar system, this effect should be just sufficiently great to be observed in the case of Mercury, which passes closest to the sun. The point of the orbit at which the planet is nearest to the sun is called the *'perihelion'*. The perihelion itself is found to be very slowly rotating around the sun, so that the orbit is as shown in Fig. 1.10. This effect is called the *'advance of the perihelion'* of Mercury. Observed and predicted values of this rotation of the planet's orbit (2×10^{-4} rad. per century) agree to within the accuracy of astronomical measurement.

3. Gravitational red shift

The third prediction of general relativity is that *a stationary clock runs more slowly in a strong gravitational field than in a weak one.*

Suppose a clock has an interval of t between ticks when it is infinitely far from all matter. When the clock is at a distance R from a mass M, then it will have an interval of

$$t = t_0/\sqrt{1-2GM/R}$$

between ticks.

An atom emitting light of a definite frequency is equivalent to a clock (atomic clock). The third prediction means that the light emitted by an atom on a star must be of frequency lower than that emitted by the same atom on earth (Since the gravitational field of star is much more stronger than that of earth). In other words, the light from the star must be shifted towards the red end of the spectrum (apart from any Doppler's shift). This effect is called the *"gravitational red shift"*. The gravitational red shift has been actually observed in the spectral lines of certain stars called *'white dwarfs'*, and also in a laboratory experiment involving *'Mossbauer effect'* and has verified the prediction.

EXERCISE

1. What is Newtonian relativity? Deduce the Galilean transformation equations.

2. Describe the Michelson-Morley experiment and explain the physical significance of negative results.

3. State and explain the basic postulates of Einstein's special theory of relativity. Derive the Lorentz space-time transformation formulae. Discuss length-contraction and time-dilation.

4. State and deduce the mathematical expression for the law of addition of relativistic velocities. Show that in no case can the resultant velocity of a material particle be greater than c.

5. Deduce the formula for relativistic variation of mass with velocity. Briefly explain its significance.

6. What is the meaning of mass-energy equivalence? Obtain Einstein's mass-energy relation. Show that $1u = 931$ MeV.

7. Write down the relativistic expression for kinetic energy of a body and show that for smaller speeds it reduces to the classical expression.

8. What is the General Theory of Relativity? Discuss the important conclusions derived from it. What are the experimental observations in favour of these conclusions?

9. Is a laboratory at rest on the earth's surface really an inertial frame reference?
[**Ans.** *No, because it is accelerated as a consequence of the earth's rotation on its axis and revolution around the sun*].

10. Obtain the volume of a cube, the proper length of each edge of which is l_0, when it is moving with a velocity v along one of its edges. [**Ans.** $l_0^3\sqrt{(1-v^2/c^2)}$]

11. Calculate the percentage contraction of a rod moving with a velocity $0.8\,c$ in a direction inclined at 60° to its own length. [**Ans.** 9%]

12. By what fraction would the diameter of the moon appear shortened to an observer in a rocket-ship travelling at a speed of 300 km/sec with respect to the moon? [**Ans.** 5×10^{-4}]

13. Determine the time (as measured by a clock at rest on the rocket) taken by a rocket to reach a distant star and return to earth with a constant velocity v equal to $\sqrt{0.9999}\ c$ if the distance to the star is 4 light years. [**Ans.** 0.08 year]

14. At what speed should a clock be moved so that it may appear to lose 1 minute in each hour?
[**Hint.** Here $t_0 = 59$ minutes, $t = 60$ minutes; $v = ?$ $v = 5.4 \times 10^7\ \text{ms}^{-1}$]

15. The half life of a particular particle, as measured in the laboratory, comes out to be 4×10^{-8} second when its speed is $0.8c$ and 3×10^{-8} second when its speed is $0.6c$. Explain this.

[**Hint.** This can be explained on the basis of relativistic time dilation.

In the first case, $t = 4 \times 10^{-8}$ sec; $v = 0.8c$

$\therefore$ Proper half life of the given particle $= t_0 = t\sqrt{(1 - v^2/c^2)} = 2.4 \times 10^{-8}$ sec;

In the second case also, $t_0 = 2.4 \times 10^{-8}$ sec. The faster it moves, the longer appears to be its half-life. Thus the variation of half-life of the given particle is due to relativistic time dilation.]

16. The length of a spaceship is measured to be exactly half its proper length. (*a*) What is the speed of spaceship relative to the observer? (*b*) What is the dilation of spaceship's unit time?
[**Ans.** (*a*) 0.866 *c*. (*b*) rate of clock will slow to half]

17. A π meson has a mean lifetime of 2×10^{-8} s when measured at rest. How far does it go before decaying into another particle if its speed is 0.99 *c*? [**Ans.** 42.1 m]

18. How fast should a rocket ship move relative to an observer in order that one year on it may correspond to two years on the earth? [**Ans.** $2.598 \times 10^8\ \text{ms}^{-1}$]

19. A rocket ship travels away from the earth at 300 ms^{-1}. How many years must elapse before a clock in the ship and one on the ground differ by one second? [**Ans.** 6.3×10^4 years]

20. In the laboratory, electrons from two accelerators are projected with the same speed of $2 \times 10^8\ \text{ms}^{-1}$ but in opposite directions. What is the relative velocity of the two sets of electrons?
[**Ans.** $2.77 \times 10^8\ \text{ms}^{-1}$]

21. Two particles come towards each other, each with speed $0.9c$ with respect to the laboratory. What is their relative speed? [**Ans.** 0.994 *c*]

22. A man on the moon observes two spaceships coming towards him from opposite directions at speed of 0.8 *c* and 0.9 *c* respectively. What is the relative speed of the two spaceships as measured by an observer on either one? [**Ans.** 0.987 *c*]

23. Calculate the mass of the electron when it is moving with a K.E. of 10 MeV.
[**Hint.** K.E. $= (m - m_0)\,c^2 = 10$ MeV $= 10 \times 1.6 \times 10^{-13}$ J

$$\therefore \quad m = m_0 + \frac{10 \times 1.6 \times 10^{-13}}{(3 \times 10^8)^2} = 9.1 \times 10^{-31} + 176 \times 10^{-31} = 185.1 \times 10^{-31}\ \text{kg.}\ ]$$

24. Compute the energy released on the conversion of 1 gram of matter into energy in kWh.
[**Ans.** 2.5×10^7 kWh].

25. What is the mass of a 1-GeV electron in terms of its rest mass? [**Ans.** 1960 m_0]

26. The total energy of a particle is exactly twice its rest energy. Calculate its speed.
[**Hint.** $mc^2 = 2m_0c^2$

$$\therefore \quad m/m_0 = 2. \text{ But } \frac{m}{m_0} = \frac{1}{\sqrt{(1 - v^2/c^2)}}; \quad \therefore \quad 2 = \frac{1}{\sqrt{(1 - v^2/c^2)}}$$

Solving, $v = 0.866\ c = 2.6 \times 10^8\ \text{ms}^{-1}$].

27. Find the speed of a 0.1 MeV electron according to classical and relativistic mechanics.
[**Ans.** $1.87 \times 10^8\ \text{ms}^{-1}$; $1.64 \times 10^8\ \text{ms}^{-1}$]

28. A particle of rest mass m_0 is moving with a velocity $0.9c$. Calculate (*a*) its relativistic mass, (*b*) its K.E. [**Ans.** (*a*) 2.3 m_0 (*b*) 1.3 m_0c^2]

29. An electron is moving at a speed of $2.7 \times 10^8\ \text{ms}^{-1}$. Find the ratio of its mass at this speed to its rest mass. Find also the total energy and the K.E. of the electron. [**Ans.** 2.29; 1.15 *MeV*; 0.65 *MeV*]

30. An electron and a positron at rest come together and annihilate each other producing two photons of equal energy. Find the energy of each photon. [**Ans.** 0.5 *MeV*]

31. If a particle could move with the velocity of light, how much K.E. would it possess? [**Ans.** ∞]

32. An astronomer on earth flashes a pulsed laser, and 1.3 sec. later the pulse reaches the moon 3.9×10^8 m away. An observer travelling in the same direction as the pulse sees the two events (*i.e.*, the flash and the arrival at the moon) as one event. What is the speed of this observer?

[**Ans.** *The observer must be moving at the speed of light in the same direction as the pulse.*]

33. Does the mass of a substance increase on melting? Why?

[**Ans.** *Yes. Because an amount of energy equal to the specific latent heat of fusion has been added to the substance.*]

34. At what velocity does the K.E. of a particle equal its rest energy? [**Ans.** 2.6×10^8 ms^{-1}]

35. Obtain an expression for the velocity of a particle in terms of its relativistic momentum and energy.

[**Hint.** $p = mv$ or $v = p/m$. But $m = E/c^2$ $\quad \therefore \quad v = pc^2/E$]

36. If photons have a speed c in one reference frame, can they be found at rest in any other frame?

[**Ans.** *No. When a body moves with a speed less than c in one reference frame, we can find another reference frame in which it is at rest. But when the body moves with speed c in one reference frame, it will move with the same speed c in all reference frames*].

37. A particle of rest mass m_0 moves with speed $c/\sqrt{2}$. Calculate its mass, momentum, total energy and kinetic energy.

[**Hint.** (*i*) $m = \dfrac{m_0}{\sqrt{1-v^2/c^2}} = \dfrac{m_0}{\sqrt{1-\dfrac{1}{2}}} = 1.41\, m_0$

(*ii*) $p = mv = \sqrt{2}\, m_0 \times c/\sqrt{2} = m_0 c$

(*iii*) $E = mc^2 = 1.41\, m_0 c^2$.

(*iv*) K.E. $= E - m_0 c^2 = 0.41\, m_0 c^2$.]

CHAPTER
2

THE PHOTON

AT A GLANCE

2.1	Particle-Wave Duality	**2.2**	Photons and Gravity
2.3	Gravitational Red Shift	**2.4**	To Show that the Rest Mass of a Photon is Zero

2.1 Particle-Wave Duality

The quantum hypothesis of Planck and the subsequent interpretation of the idea by Einstein gave electromagnetic radiation discrete properties somewhat similar to those of a particle. These quantized components of light became known as *photons.* According to Einstein, light of frequency ν consists of a shower of corpuscles or photons each of energy $h\nu$. It was surprising that such discrete characteristics should be in evidence since light was firmly established as a wave phenomenon. The quantum theory of light which treats it strictly as a particle phenomenon, incorporates the light frequency ν, a wave concept. The quantum theory made provision for radiation to have *both* wave and particle aspects in a complementary form of coexistence.

Let us discuss the connection between the wave theory of light and the quantum theory of light.

(*i*) The radiation field is a continuous system of propagating electric and magnetic oscillations. We choose the electric field **E** to specify the form of the propagating wave. Consider a monochromatic wave of frequency ν that is incident on a screen. The flow of energy per unit time across unit area is the intensity of the wave I.

$$I = \varepsilon_0 c\overline{E}^2 \qquad \text{...(1)}$$

Here $\overline{E}^2$ is the average of the square of the instantaneous magnitude of the wave's electric field over a complete cycle.

(*ii*) Consider the photon model of the same wave. The intensity of radiation (I) is equal to the number of photons (N) crossing unit area per second multiplied by the energy ($h\nu$) of the photon *i.e.*,

$$I = N h\nu \qquad \text{...(2)}$$

Both descriptions must give the same value for I.

$$\therefore \qquad N\,h\nu = \varepsilon_0 c \overline{E}^2 \qquad \text{...(3)}$$

This formula is a statement of particle-wave duality since the left side is written in the language of discrete quanta while the right side is written in the language of wave fields.

The rate of arrival of photons at the screen is given by

$$N = \frac{\varepsilon_0 c}{h\nu} \overline{E}^2 \qquad \text{...(4)}$$

If N is large, an observer looking at the screen would find a continuous distribution of light whose pattern corresponds to the distribution of $\overline{E}^2$. This shows the validity of the wave theory of light. If N is small, the observer would find a series of apparently random flashes. This points to light being a quantum phenomenon.

Consider an experimental arrangement in which light is diffracted by a double slit and is detected on a screen (Fig. 2.1). Suppose the light beam is very weak and there is only one photon at a time in the apparatus. We can regard each photon as having a wave associated with it. The two slits transmit the photon in the manner of a wave passing through *both* slits at once. When it strikes the screen, light is behaving as a particle does. Thus the wave and particle pictures give complementary descriptions of the same system. Bohr introduced the **complementarity principle** *which states that* the wave and particle aspects of matter are complementary rather *than being contradictory, both equally essential for a full description of the phenomenon.*

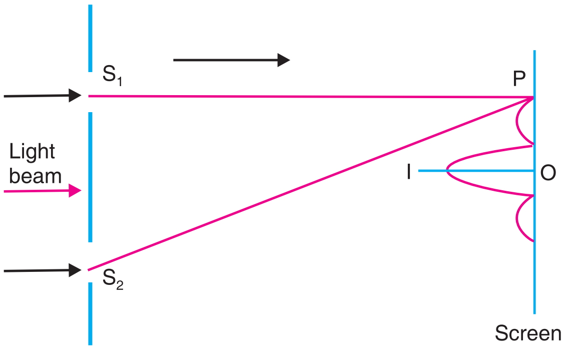
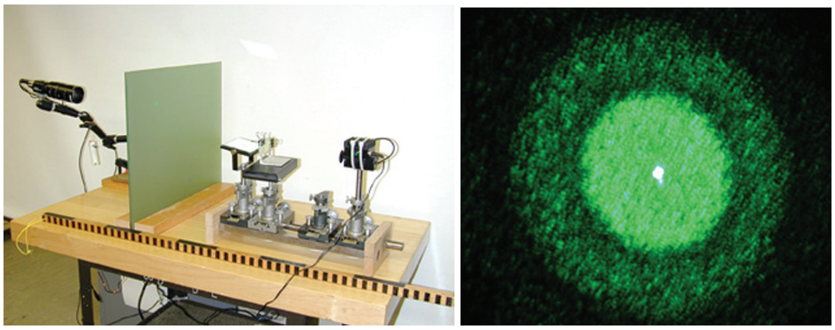


Fig. 2.1



Experimental Setup.

2.2 Photons and Gravity

The second postulate of general relativity states that *there is no way for an observer in a closed laboratory to distinguish between the effects produced by a gravitational field and those produced by an acceleration of the laboratory.* This postulate is known as the *principle of equivalence.* This postulate follows from the experimental observation that the inertial mass of a body is always exactly equal to its gravitational mass.

For instance, if a man is inside a closed elevator car in outer space (Fig. 2.2), and if the car is suddenly accelerated upwards, he cannot decide whether a gravitational force had pulled him downwards or the car had been accelerated upwards.

The most interesting prediction of the general theory of relativity is that light passing near a large mass will be bent. The following two lines of reasoning lead to the conclusion that light should be affected by gravity.

1. The existence of momentum for a photon implies that it must have an effective mass *m* also. This mass can be computed by mass-energy relation ($E = mc^2$).

$$m = \frac{E}{c^2} = \frac{h\nu}{c^2}.$$

Because gravitational mass is inseparable from inertial mass, light ought to be subject to gravity.

2. Consider a spaceship with a hole in the side allowing a ray of light to enter. If it is accelerating upward as shown in fig. 2.3 (*a*), an observer inside will say that the light is bent down in a parabolic path as shown. The bending of the light is due to the motion of the rocket ship during the period of time it takes the light to travel from the entrance hole to the far side. The principle of equivalence indicates that exactly the same observation should be made in a rocket ship at rest on the surface of a large planet [Fig. 2.3 (*b*)]. The curvature of space-time will cause a light ray to be bent in the direction of the acceleration due to gravity.

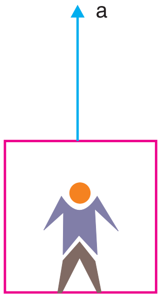
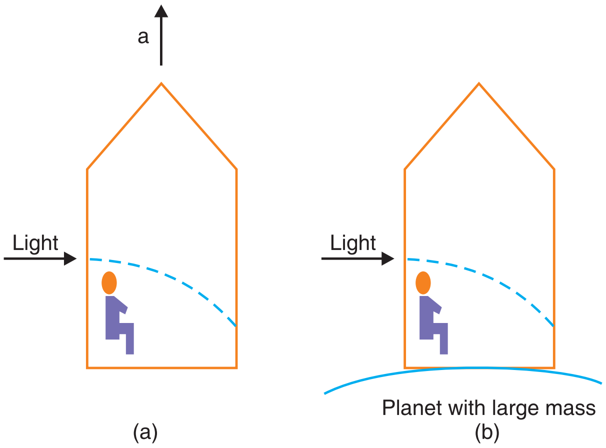


Fig. 2.2

The bending of light in a gravitational field is observed in the deflection of light from stars, moving past the sun. During a solar eclipse a star *S* (Fig. 2.4) near the sun appears to be displaced to *S*′. This is because the light from the star passing near the sun is bent in sun's gravitational field. The experimental value of the observed deflection is almost equal to its theoretical value which is 1.75 seconds. Thus *the theoretical prediction of the bending of light rays in the gravitational field agrees with the observations.*



(a)

(b)

Accelerated Rocketship

Rocketship at rest in gravitational field

Fig. 2.3

The gravitational behaviour of light can also be demonstrated in the laboratory. Consider a photon falling through a height *H*. The photon gains the energy *mgH*. This gain in energy is manifested as an increase in frequency from ν to ν′.

Final photon energy ($h\nu'$) = Initial photon energy ($h\nu$) + increase in energy (mgH)

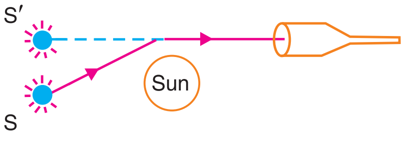


Fig. 2.4

$$h\nu' = h\nu + \frac{h\nu}{c^2} gH$$

This result has been verified by Pound and Rebka (1960) using Mossbauer effect.

2.3 Gravitational Red Shift

The general theory of relativity predicts that the frequency of a spectral line due to an atomic transition will appear to be less if the light travels against the gravitational field before being measured by an observer at some distance from the source emitting the light. Hence the measured wavelength will shift towards the red. Let us derive an expression for gravitational red shift.

Consider a star of mass M and radius R (Fig. 2.5). Suppose a photon of initial frequency ν is emitted by the star.

Mass of the photon $= m = h\nu/c^2$.

The potential energy of the photon on the star's surface is

$$V = -\frac{GMm}{R} = -\frac{GMh\nu}{c^2R}$$

Quantum energy of the photon $= h\nu$

The total energy E of the photon is

$$E = h\nu - \frac{GMh\nu}{c^2R} = h\nu\left(1 - \frac{GM}{c^2R}\right)$$

When the photon arrives at the earth, it is beyond the star's gravitational field but its total energy remains the same. The photon's energy is now entirely electromagnetic. Let ν' be the frequency of the photon at the earth. Then,

$$E = h\nu'$$

$$\therefore \quad h\nu' = h\nu\left(1 - \frac{GM}{c^2R}\right)$$

$$\frac{\nu'}{\nu} = 1 - \frac{GM}{c^2R}$$

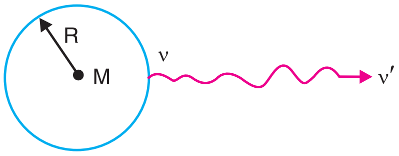


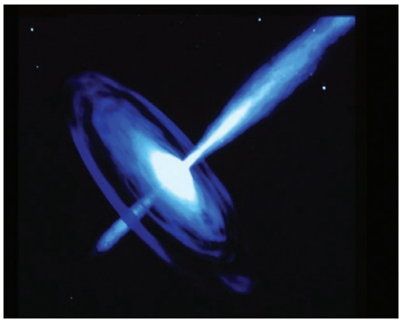
Fig. 2.5

The relative frequency change is

$$\frac{\Delta\nu}{\nu} = \frac{\nu - \nu'}{\nu} = 1 - \frac{\nu'}{\nu} = \frac{GM}{c^2R} \qquad \text{...(1)}$$

Thus there is a red shift for the spectral lines of light emitted by an atom situated in a gravitational field when this light is observed in a gravitation free part of space.

Black Hole. Another prediction of the general theory that was pointed out in 1939 by Oppenheimer and Snyder was the possibility of what is called a *black hole*. Consider a star so dense that $GM/(c^2R) \geq 1$. Then from Eq. (1) we see that no photon can ever leave the star. Such a star cannot radiate and so would be invisible—a *black hole* in space.



Massive Black Hole.

Now the question is how large would M have to be and how small would R have to be in order that the escape velocity would be equal to c? According to the general theory, if an object of mass M has a radius given by

$$R_s = \frac{2GM}{c^2} \qquad \text{...(2)}$$

the escape velocity would be c. This expression was derived by Schwarzschild. R_s is called *Schwarschild radius.*

Since a black hole cannot obviously be seen in the usual sense, evidence for its existence can only be indicated by the effect on matter in its vicinity. A black hole that is a member of a double-star system will manifest its presence by its gravitational effect on the other member. Another possibility would be to observe x radiation given off by charged particles as they are accelerated toward a black hole.

2.4 To Show that the Rest Mass of a Photon is Zero

A photon travels with the velocity of light. Hence we must use the relativistic expression for its momentum, p.

Thus, for a phton

$$p = \frac{m_0 v}{\sqrt{[1-(v^2/c^2)]}}, \qquad \text{... (1)}$$

Here, m_0 is the rest mass and v the velocity of the photon

From quantum theory of radiation, the momentum of a photon of radiation of wavelength λ is

$$p = \mathrm{h}/\lambda, \qquad \text{... (2)}$$

Here, h is the Planck's constant.

$$\therefore \quad \frac{h}{\lambda} = \frac{m_0 v}{\sqrt{[1-(v^2/c^2)]}} \qquad \text{... (3)}$$

or

$$m_0 = \frac{h}{v\lambda}\sqrt{[1-(v^2/c^2)]}.$$

Since for photon, $v = c$, we have

$$m_0 = 0.$$

Hence the rest mass of photon is zero.

Example 1. *Calculate the Schwarzschild radius of a black hole of mass 15 × 10^6 M_O.*

Sol. $M = 15 \times 10^6\ \mathrm{M_O} = 15 \times 10^6 \times$ (sun's mass) $= 15 \times 10^6 \times 1.9599 \times 10^{30}$ kg,

$G = 6.673 \times 10^{-11}$ Nm^2 kg^{-2}, $c = 3 \times 10^8$ ms^{-1}; *Schwarzschild radius*, R_s = ?

$$R_s = \frac{2GM}{c^2} = \frac{2\times 6.6673\times 10^{-11}\times 15\times 10^6\times 1.9599\times 10^{30}}{(3\times 10^8)^2} = 4.425 \times 10^7 \text{ m}$$

Example 2. *The sun's mass is 2 × 10^{30} kg and its radius is 7 × 10^8 m. Find the gravitational red shift in light of wavelength 500 nm emitted by the sun.*

Sol. The relative frequency change is

$$\frac{\Delta\nu}{\nu} = \frac{GM}{c^2 R} \text{ or } \Delta\nu = \nu\cdot\frac{GM}{c^2 R}$$

$$\Delta\left(\frac{c}{\lambda}\right) = \frac{c}{\lambda}\frac{GM}{c^2 R} \text{ or } -\frac{c}{\lambda^2}d\lambda = \frac{c}{\lambda}\cdot\frac{GM}{c^2 R} \text{or } |d\lambda| = \lambda\cdot\frac{GM}{c^2 R}$$

$$\therefore \quad d\lambda = \frac{500\times 10^{-9}\times 6.7\times 10^{-11}\times 2\times 10^{30}}{(3\times 10^8)^2\times 7\times 10^8}\text{m} = 1.06 \times 10^{-12} \text{ m.}$$

CHAPTER 3

THEORY OF RELATIVITY II

AT A GLANCE			
3.1	Space-Time Diagrams	**3.2**	Geometrical Representation of Simultaneity, Contraction and Dilation

3.1 Space-Time Diagrams

For simplicity we shall consider only one space-axis, the X-axis and shall ignore the Y and Z axes. The co-ordinates of an event are then represented by (x, t). All possible space-time co-ordinates can be represented on a space-time diagram in which the space axis is horizontal and the time-axis is vertical. For convenience the dimensions of the co-ordinates are kept the same; this is easily done by using $m = ct$. Then the Lorentz transformation equations for x and t can be written as follows:

$$\left.\begin{aligned} x' &= \frac{x-\beta m}{\sqrt{(1-\beta^2)}},\ x = \frac{x'+\beta m'}{\sqrt{(1-\beta^2)}} \\ m' &= \frac{m-\beta x}{\sqrt{1-\beta^2}},\ m = \frac{m'+\beta x'}{\sqrt{(1-\beta^2)}} \end{aligned}\right\} \qquad \text{...(1)}$$

We draw the x and m axes of frame S orthogonal to one another (Fig. 3.1). Then the motion

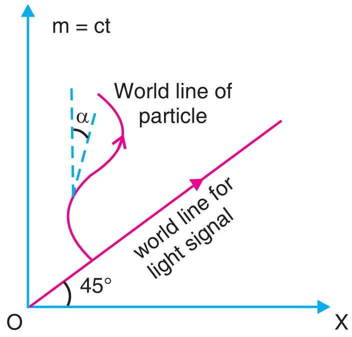


Fig. 3.1

of a particle in this frame is represented by a curve known as *world line*, which gives the loci of space-time points, corresponding to the motion. The tangent to the world line at any point is always inclined at an angle less than 45° with the time axis, since this angle is given by

$$\tan\alpha = \frac{1}{c}\frac{dx}{dt} = \frac{u}{c}$$

and we must have $u < c$ for a material particle. The world line for light signal (*i.e.*, $u = c$) is a straight line making an angle 45° with the axes.

3.2 Geometrical Representation of Simultaneity, Contraction and Dilation

Consider two systems S and S'. System S' is moving with velocity v relative to system S along positive direction of X-axis.

Simultaneity: The two events will be simultaneous in system S' if they have the same time co-ordinate m'. Hence if the events lie parallel to X'-axis, they are simultaneous in system S'. For example, events P_1 and P_2 in Fig. 3.2 are simultaneous in system S'. But they are not simultaneous in system S, because they have different time co-ordinates m_1 and m_2 in system S. Similarly, the two events Q_1 and Q_2 which are simultaneous in system S, are not simultaneous in system S', because they have different time co-ordinates m_1' and m_2' in system S'.

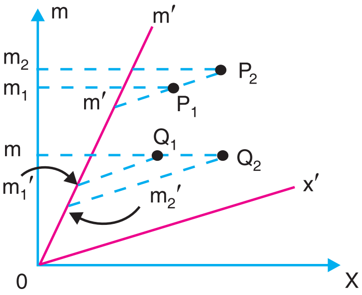


Fig. 3.2

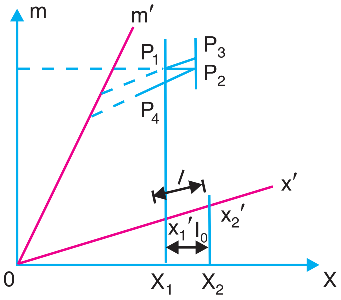


Fig. 3.3

Space Contraction: Consider a rod of rest length l_0 at rest in system S (Fig. 3.3). Its end points are at $x = x_1$ and $x = x_2$. As time goes on, the world line of each end point traces out a vertical line parallel to m-axis. By definition, the length of the rod is the distance between its end points measured simultaneously (*i.e.*, $l_0 = x_2 - x_1$). In the rest system S, the length of the rod is the distance in system S between the intersections of the world lines with the X-axis or any line parallel to X-axis, because these intersecting points represent simultaneous events in system S. The rest length is l_0. To find the length of the rod in system S', relative to which the rod moves, we must find the distance in frame S' between end-points measured simultaneously. This will be the separation in system S' of the intersections of the world lines with X' - axis or any line parallel to X ‘ - axis, because these intersecting points represent simultaneous events in system S'. The length l of the rod in system S' is clearly less than l_0.

From Fig. 3.3, it is clear that different lengths in different frames are due to relativity of simultaneity. In fact the two observers do not measure the same pair of events in determining the length of the rod (*e.g.*, the S observer uses P_1 and P_2 ; But the S' observer uses P_1 and P_3 or P_2 and P_4), because the events which are simultaneous to one inertial observer are *not* simultaneous to the other.

Time-Dilation: To demonstrate the time-dilation result geometrically, consider Fig. 3.4. A clock at rest in system S ticks off units of time there. Let T_1 and T_2 be the events of ticking at $m = 2$ and $m = 3$. The time interval in system S between ticks is unity. Here solid vertical line T_1T_2 is the world line corresponding to the clock under consideration in system S. This clock is in motion relative to S'. So the clock is at a different place there each time it ticks. So to measure the time interval between events T_1 and T_2 in S', two different clocks, one at the location of event T_1 and the other at the location of event T_2 must be used. The difference in reading of these clocks in S' is the difference in times between events T_1 and T_2 as measured in S'. From the figure, we see that this time interval is greater than unity. Hence relative to system S' the clock at rest in system S appears to run slow, *i.e.*, the moving clock appears to run slow.

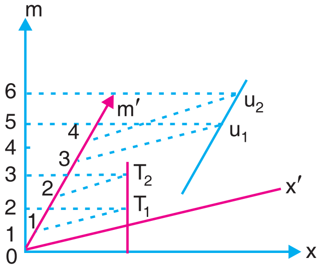


Fig. 3.4

PART – II

ATOMIC PHYSICS

4. The Electron and Band Theory of Solids
5. Positive Rays and Particle Properties of Waves
6. Structure of the Atom
7. X-Rays
8. The Photoelectric Effect
9. Planck's Quantum Theory
10. Electron

TECHNOLOGY
• Immersion Gold
• Buried Capacitance
• PCMCIA
• BGA
• Controlled Impedance
• Fine Pitch <16 mil
• All Processes In-House
• Entek

CHAPTER

4

THE ELECTRON AND BAND THEORY OF SOLIDS

AT A GLANCE

4.1	Determination of the Electronic Charge : Millikan's Oil-drop Method	**4.2**	The Free Electron Theory of Metals
4.3	Expression for Electrical Conductivity	**4.4**	Expression for Thermal Conductivity
4.5	Electron Microscope	**4.6**	Band Theory of Solids
4.7.	Classification of Solids on the Basis of Band Theory	**4.8**	Optical Properties of Solids
4.9	Energy Bands: Alternative Analysis		

4.1 Determination of the Election Charge : Millikan's oil-drop Method

Experimental arrangement. The apparatus consists of two optically plane parallel brass plates *A* and *B* (Fig. 4.1), separated by insulating rods of glass. The lower plate *B* is earthed. The upper plate *A* can be charged to a positive potential of the order of 10000 volts, from a high-tension battery. This observation chamber is situated in a bigger chamber *C* which is completely surrounded by a constant temperature bath of oil *E*. Tiny drops of a heavy non-volatile oil are sprayed into the chamber by an atomizer *D*. Some of the oil drops enter the space between the plates through the pinhole *H* in the top plate *A*. These drops are charged, due to the frictional effect at the nozzle of the atomizer. The air between the plates can be ionized by allowing X-rays to pass through it. Then the drops may pick up additional charges. The drops are illuminated by light from an arc lamp *S*. The oil-drop is observed by means of a short-focus telescope *T* provided with a millimetre scale in the eyepiece.

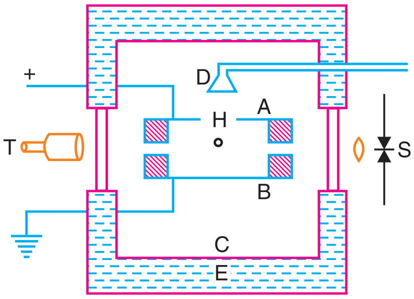
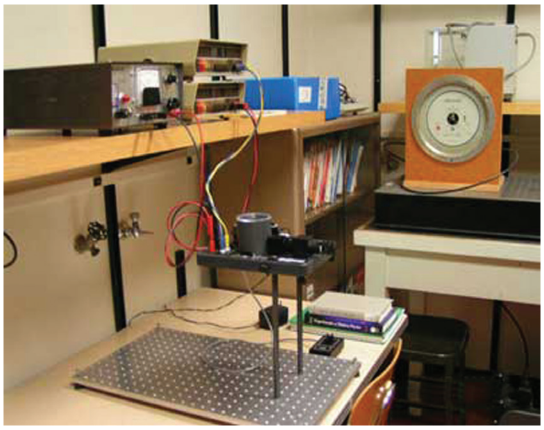


Fig. 4.1

Experimental Procedure. The plates are first connected together, so that the electric field between them is zero. A single drop is observed through the microscope as it descends under the action of gravity. As the drop is moving in a viscous medium (air), it will move with a constant terminal velocity v. Using a stop watch, the terminal velocity of the drop is measured. Let

v = terminal velocity of the free fall,
a = radius of the drop,
ρ = density of oil,
σ = density of air,
η = coefficient of viscosity of air,
m = effective mass of the drop and
g = acceleration due to gravity.



Millikan Oil Drop Experiment (Charge of an Electron).

Then by Stokes law, $mg = 6\pi\eta a v$ ...(1)

$$\text{or}\quad \frac{4}{3}\pi a^3(\rho-\sigma)g = 6\pi\eta a v \qquad \text{...(2)}$$

$$\text{or}\quad a = \left[\frac{9\eta v}{2(\rho-\sigma)g}\right]^{\frac{1}{2}} \qquad \text{...(3)}$$

$$\therefore \text{ The effective mass of the drop} = m = \frac{4}{3}\pi a^3(\rho-\sigma) = \frac{4}{3}\pi\left[\frac{9\eta v}{2(\rho-\sigma)g}\right]^{\frac{3}{2}}(\rho-\sigma) \qquad \text{...(4)}$$

Thus, by measuring v, we can calculate the effective mass m of the chosen oil drop.

Now, an electrostatic field X is applied steadily, so that the drop moves upwards with a steady velocity v_1. If the total charge on the drop is $n_1 e$ (where n_1 is an integer),

the resultant upward force on the drop = $Xn_1e - mg$.

According to Stokes law, $Xn_1e - mg = 6\pi\eta a v_1$ ...(5)

$$\text{Dividing (5) by (1),}\quad \frac{Xn_1 e - mg}{mg} = \frac{v_1}{v} \quad\text{or}\quad \frac{Xn_1 e}{mg} = \frac{v_1+v}{v}$$

$$\therefore \qquad n_1 e = \frac{mg(v_1+v)}{Xv} \qquad \text{...(6)}$$

Now, the air between the plates is irradiated with X–rays. X–rays ionize the air molecules. The oil drop under observation may pick up one or more ionic charges. Due to the acquisition of charges, the velocity of the drop changes suddenly, when X is constant. Let the new velocity of the drop be v_2 and the corresponding charge on the drop $n_2 e$. Then,

$$n_2 e = \frac{mg(v_2+v)}{Xv} \qquad \text{...(7)}$$

$$\therefore \qquad (n_2-n_1)e = \frac{mg}{Xv}(v_2-v_1)$$

If $v_2 > v_1$, $(n_2-n_1)e$ is positive and if $v_2 < v_1$, $(n_2-n_1)e$ is negative. Both are possible, as both positive and negative ions are present.

Now for the same drop under the same electric field (X), mg/Xv is constant. Hence $v_2 - v_1 \propto (n_2 - n_1)\, e$. *i.e.*, any change in the velocity of the drop is proportional to the quantity of the acquired charge. Millikan made thousands of observations on a single drop. He found that there was a minimum value of $(v_2 - v_1)$ and that other values of $(v_2 - v_1)$ were simple integral multiples of this $(v_2 - v_1)_{minm}$. $(v_2 - v_1)_{minm}$ corresponds to the addition or subtraction of one unit of charge. *i.e.*, $(n_2 - n_1) = 1$ when $(v_2 - v_1)$ is minimum.

$$\therefore \qquad e = \pm \frac{mg}{Xv}(v_2 - v_1)_{minm} \qquad ...(8)$$

Substituting the value of m from (4),

$$e = \pm \frac{4}{3}\pi \left(\frac{9\eta}{2}\right)^{\frac{3}{2}} \left[\frac{v}{(\rho - \sigma)\, g}\right]^{\frac{1}{2}} \frac{(v_2 - v_1)_{minm}}{X} \qquad ...(9)$$

The value of electronic charge e found by Millikan was 1.591×10^{-19} C. The value of e now generally accepted is 1.602×10^{-19} C.

Balanced Oil-Drop Method. In this method, the intensity of the electric field is adjusted to a suitable value X such that the drop remains at rest. Then the effective weight of the drop (mg) exactly balances the force due to the electric field (Xe).

$\therefore$ $Xe = mg$ or $e = mg/X$. (We have assumed that the drop carries one charge).

Substituting the value of m from (4),

$$e = \frac{4}{3}\frac{\pi}{X}\left(\frac{9\eta v}{2}\right)^{\frac{3}{2}} \frac{1}{[(\rho - \sigma)\, g]^{1/2}}$$

4.2 The Free Electron Theory of Metals

Electrical Conduction in Metals. The following assumptions are made in the electron theory:

(*i*) In *conductors* there is a large number of *free electrons,* actually separated from the atoms. They are capable of moving freely in the interatomic spaces.

(*ii*) These electrons (conduction electrons) are treated as an *electron gas*. The electrons are performing random motion with all possible velocities like the gas molecules. The mean K.E. of the electrons is equal to that of the gas molecules at the same temperature. The positive ions in the metal occupy fixed positions.

In the absence of an electric field, the free electrons move about irregularly in the whole volume of the metal like the gas molecules. When an electric field is applied between the ends of a conductor, there will be an electron drift in a direction opposite to the direction of the electric field, since the electrons carry a negative charge. The effect of the electric field is to impart a velocity called "*drift velocity*" to the electron in addition to the velocity of random motion.

4.3 Expression for Electrical Conductivity

Let m = mass of the electron ;

v = mean velocity of the electron corresponding to temperature equilibrium with the metal and

T = absolute temperature of the metal.

According to kinetic theory, $\frac{1}{2} mv^2 = aT$ ...(1)

where a is a constant

When an electric field X is applied between the ends of a conductor, each free electron experiences a force Xe. The acceleration of the electron in the direction of the field is Xe/m. If λ is the mean free path of the electron, then the time t between two consecutive collisions is given by

$$t = \lambda/v \qquad \text{...(2)}$$

The average distance travelled by the electron in the direction of the field during the time t

$$= s = \frac{1}{2}\left(\frac{Xe}{m}\right)\frac{\lambda^2}{v^2} \qquad \text{...(3)}$$

The average drift velocity V of the free electrons parallel to the direction of the field is given by

$$V = \frac{s}{t} = \frac{Xe\lambda^2}{2mv^2} \times \frac{v}{\lambda} = \frac{Xe\lambda}{2mv} \qquad \text{...(4)}$$

Substituting the value of m from equation (1),

$$V = \frac{Xe\lambda v}{4aT} \qquad \text{...(5)}$$

Let n be the number of electrons per unit volume of the metal.

Then the current density in the metal $= J = neV$ ...(6)

Substituting the value of V from equation (5),

$$J = \frac{nXe^2\lambda v}{4aT} \qquad \text{...(7)}$$

Electrical conductivity of the metal $= \sigma = \dfrac{\text{Current density}}{\text{Electric field}}$

$$\therefore \qquad \sigma = \frac{J}{X} = \frac{ne^2\lambda v}{4aT} \qquad \text{...(8)}$$

n is different for different metals. Hence the difference in the conductivity of different metals is due to the difference in the number of free electrons in them.

The specific resistance of the metal $= \rho = \dfrac{1}{\sigma} = \dfrac{4aT}{ne^2\lambda v}$ ...(9)

Equation (9) shows that the specific resistance of the metal (ρ) is directly proportional to the absolute temperature T.

Limitations of the Theory. (1) It does not give a satisfactory explanation as to why certain metals like copper and silver have high conductivity.

(2) It is unable to explain adequately the variation of resistance of pure metals at the temperature of liquid helium.

4.4 Expression for Thermal Conductivity

The conduction of heat in a metal is also mostly through free electrons. Consider any plane *AB* in a conductor normal to the direction of flow of heat (Fig. 4.2). Consider two parallel planes *CD* and *EF*, on either side of *AB* at a distance equal to the mean free path of the electrons λ. Let the temperatures of the planes *CD* and *EF* be T_1 and T_2 ($T_1 > T_2$). The distance between the two planes = 2λ. The K.E. of the molecules at the hotter end is greater than that of the molecules at the colder end. As result of collisions, the molecules at the hot end lose K.E. while those at the cold end gain K.E. The conduction of heat through the element is due to this transference of energy.

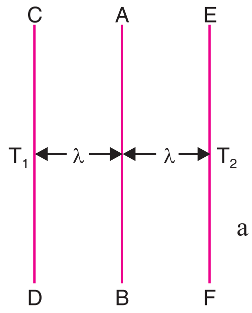


Fig. 4.2

Let v_1 and v_2 be the velocities of the electron corresponding to the temperatures T_1 and T_2. Then

$$\frac{1}{2}mv_1^2 = aT_1 \text{ and } \frac{1}{2}mv_2^2 = aT_2.$$

If v is the mean velocity of each electron while crossing AB, the number of electrons crossing unit area of section AB per second is $(n/6)\,v$. (n = the number of electrons in unit volume).

$$\left.\begin{array}{l}\text{Heat energy transferred across}\\ \text{unit area of the plane } AB \text{ per}\\ \text{second from left to right}\end{array}\right\} = \frac{n}{6}v \times \frac{1}{2}mv_1^2 = \frac{n}{6}vaT_1.$$

Similarly,

$$\left.\begin{array}{l}\text{Heat energy transferred across}\\ \text{unit area of the plane } AB \text{ per}\\ \text{second from right to left}\end{array}\right\} = \frac{n}{6}\,vaT_2.$$

$$\left.\begin{array}{l}\therefore\ \text{Net heat energy transferred}\\ \text{per unit area of the plane } AB\\ \text{per second}\end{array}\right\} = \frac{n}{6}\,va\,(T_1 - T_2)$$

temperature gradient = $(T_1 - T_2) / 2\lambda$.

The coefficient of thermal conductivity (K) is the ratio of the quantity of heat transferred across unit area in one second to the temperature gradient.

$$\therefore \qquad K = \frac{(n/6)\,va\,(T_1 - T_2)}{(T_1 - T_2)/2\lambda} = \frac{1}{3}\,nva\lambda$$

Wiedemann-Franz Law: Statement. *The ratio of thermal conductivity to electrical conductivity is a constant for all metals and is proportional to the absolute temperature.*

Explanation.

$$\frac{K}{\sigma} = \frac{\frac{1}{3}\,nva\lambda}{\left(\frac{ne^2\lambda v}{4aT}\right)} = \frac{4}{3}\frac{a^2}{e^2}T.$$

Since $\frac{4}{3}\frac{a^2}{e^2}$ = a constant, $\frac{K}{\sigma} \propto T$.

This relation shows that the ratio K/σ is independent of the nature of the metal and is proportional to the absolute temperature.

Thus, $\frac{K}{\sigma} \propto T$ or $\frac{K}{\sigma T} = \frac{4}{3}\frac{a^2}{e^2}$. This quantity is known as Lorentz's number.

Among the metals, the best electrical conductors (*Ag*, *Cu*, *Al*, *Au*) are also the best heat conductors.

4.5 Electron Microscope

Principle. A particle of mass m moving with a large velocity v behaves like a wave of wavelength $\lambda = h/mv$. This wave nature of the particle is used in the construction of an electron microscope. An electron microscope is, in principle, similar to that of an optical microscope. In optical microscope, using visible light, focusing is done by suitable lenses. But in an electron microscope, a beam of electrons is used in place of ordinary light and focusing is done by electric and magnetic fields.

The wave theory of matter shows, that electrons accelerated to a potential of 60000 volt have a wavelength of about 5×10^{-12} m. This wavelength is 10^5 times shorter than the wavelength for visible light. As the resolving power of a microscope is proportional to the wavelength of light used, the resolving power of an electron microscope is much higher than that of an optical microscope.

Electrostatic lens. An Electron beam passing through an electric field undergoes divergence or convergence, depending upon the potential gradients at different points in the field. The arrangement shown in Fig. 4.3 works like a lens for a beam of electrons. Metallic cylinder *A* is connected to the negative terminal of a source of high potential difference and *B* to its positive terminal. The electric lines of force between them are almost semicircular in shape. A beam of electrons moving coaxially to *A*, gets converged, when it passes through the electrostatic field between *A* and *B*. This system of cylinders is called an electrostatic lens and its focal length depends on the potential difference between the cylinders.

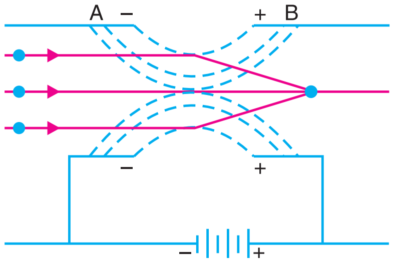
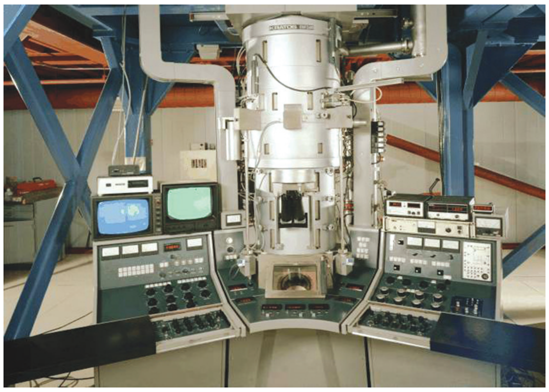


Fig. 4.3



1.5 MeV Electron Microscope.

The focal length of the electrostatic lens can be changed by adjusting the P.D. of the cylinders. But the focal length of a glass lens is fixed, depending upon its radius of curvature and its material.

Magnetic lens. In a magnetic field, an electron beam undergoes divergence or convergence depending upon the potential gradient of different points in the field. An arrangement of magnetic coils used for converging or diverging an electron beam is called a *magnetic lens.* The *focusing* action of a magnetic field on electrons is identical with that of an optical lens on a beam of light and hence the arrangement is called *magnetic lens.*

Fig. 4.4 shows the electrons being emitted by a source *O*, then passing across a magnetic field and getting converged to a point *I*. The magnetic field is generally provided by coils carrying current. Two coils C_1 and C_2 are enclosed in two iron cases which have one hole each. When the holes face each other, the magnetic field in the space between the two produces convergence. The focal length of such a magnetic lens depends upon the quality of the coil and the current it carries. Adjusting the currents in the coils, the intensity of the magnetic field can be adjusted and thus the focal length of the system can be adjusted. In electron microscope generally magnetic lenses are used.

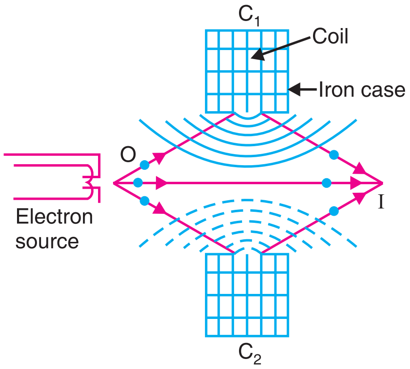


Fig. 4.4

Construction. The essential parts of an electron microscope are shown in Fig. 4.5. An optical microscope is shown alongside for comparison. Electrons liberated from the source are accelerated under potentials of about 60000 V. The accelerated electron beam is then focused by a magnetic condensing lens on the object. The electrons which are scattered by the object pass on, go through the magnetic objective lens which forms the first magnified image. The projection lens forms a real magnified image on a fluorescent screen.

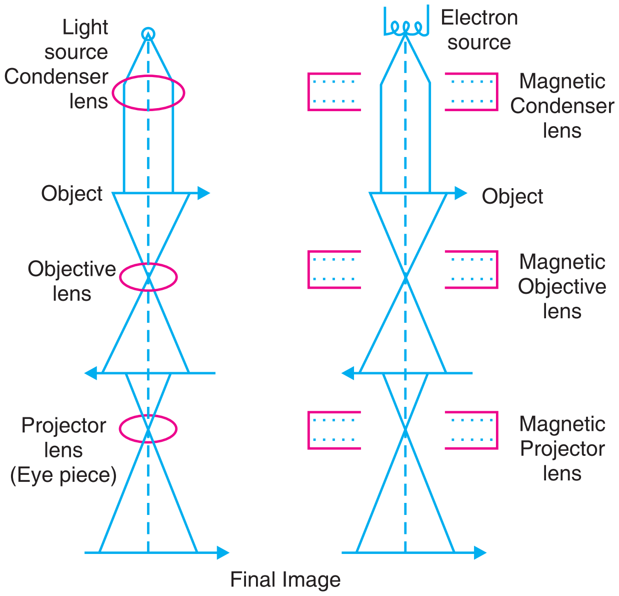

Light
source
Condenser
lens
Object
Objective
lens
Projector
lens
(Eye piece)
Electron
source
Magnetic
Condenser
lens
Object
Magnetic
Objective
lens
Magnetic
Projector
lens
Final Image

Fig. 4.5

Uses. (1) It is used in industry to study the structure of textile fibres, surface of metals, composition of paper and paints etc.

(2) In medicine and biology it is used to study virus, the disease causing agent beyond the reach of ordinary microscope. The *bacteria* are shown in greater detail in an electron microscope. A knowledge of the structure of viruses enables scientist to devise methods of their destruction and thereby control diseases.

(3) In Physics, it has been used in the investigation of atomic structure and structure of crystals.

(4) In Chemistry, small particles forming colloids have become open to study and analysis.

4.6 Band Theory of Solids

A solid contains an enormous number of atoms packed closely together. Each atom, when isolated, has a discrete set of electron energy levels, 1*s*, 2*s*, 2*p*, ... If we imagine all the *N* atoms of the solid to be isolated from one another, they would have completely coinciding schemes of their energy levels. The energies of electrons within any one isolated atom obey the following conditions.

(*i*) There are specific electronic energy levels around each atom [(Fig. 4.6 (*a*)]. Electrons cannot occupy spaces between these levels.

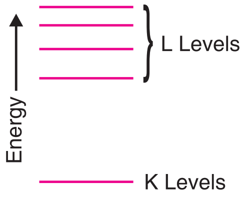

Energy
L Levels
K Levels

Fig. 4.6 (a)

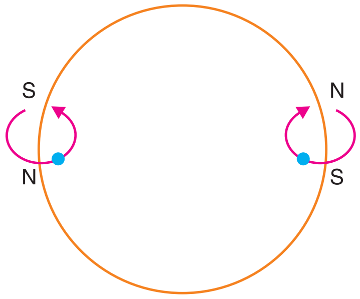


Fig. 4.6 (b)

(*ii*) Electrons fill the lowest energy levels first. A specific quantity of energy, called a *quantum of energy*, must be supplied to move an electron to the next higher level.

(*iii*) According to the Pauli exclusion principle, no two electrons can occupy the same quantum state. Not more than two electrons may occupy any one energy level. Two electrons may occupy the same energy level because they have opposite magnetic characteristics *i.e.*, *electron spins* are in opposite directions [(Fig. 4.6 (*b*)].

Now, let us look into what happens to the energy levels of isolated atoms, as they are brought closer and closer together, to form a solid. If the atoms are brought in close proximity, the valence electrons of adjacent atoms interact. Hence the valence electrons constitute a single system of electrons common to the entire crystal, their outermost electronic orbits overlapping. Therefore, the N electrons will now have to occupy different energy levels. This is brought about by the electric forces exerted on each electron by all the N nuclei. As a result of these forces, each atomic energy level is split up into a large number of closely spaced energy levels. A set of such closely spaced energy levels is called an *energy band*.

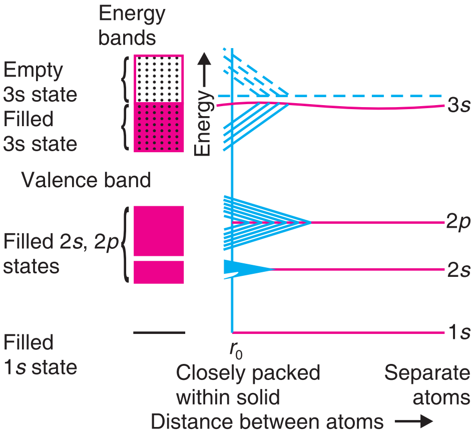


Fig. 4.7

As an example, the 11 electrons in a neutral sodium atom, each occupy a specific energy level as indicated in Fig. 4.7. The energy levels of sodium become bands when the atoms are close together. In the figure, r_0 represents the spacing between atoms in solid sodium. When the atoms are part of a solid, they interact with each other and the electrons have slightly different energies.

In an energy band, the allowed energies are almost continuous. These energy bands are separated by regions, which have no allowed energy levels. These regions are known as "*forbidden bands*" or "*energy gaps*" [Fig. 4.8].

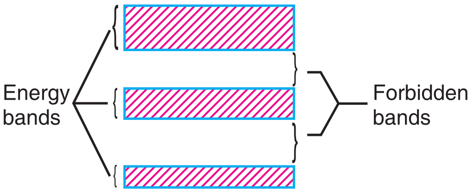


Fig. 4.8

The amount of splitting is not the same for different levels. The levels filled by the valence electrons in an atom are disturbed to a greater extent, while those filled by inner electrons are disturbed only slightly. Fig. 4.7 shows the splitting of the levels as a function of the distance r between the atoms.

If there are N atoms in a solid, there will be N allowed quantum states in each band. Each quantum state can be occupied by a maximum of two electrons with opposite spins. Thus each energy band can be occupied by $2N$ electrons.

The *valence band* consists of a group of states containing the outermost or the *valence* electrons of an atom. The band formed from the atomic energy levels containing valence electrons is called 'valence band'. These electrons have the highest energy. The band is, obviously, the highest occupied band. Above the valence band, there is the band of next higher permitted energies called the *conduction band.* Conduction band is separated by a *gap* from the valence band (Fig. 4.9). The interval between them represents the energies which their electrons cannot possess. The

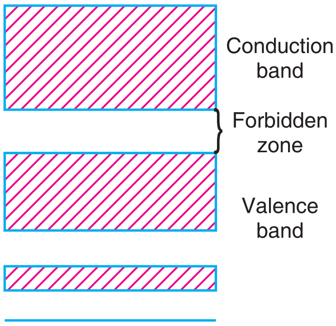


Fig. 4.9

conduction band corresponds to the first excited states and is normally the lowest *unfilled* energy band. In this band, the electrons can move freely and the electrons are generally called "conduction electrons".

4.7 Classification of Solids on the Basis of Band Theory

Whether a solid is an insulator, conductor or semiconductor, is determined by the energy band structure, the arrangement of electrons with respect to these bands and the width of forbidden bands.

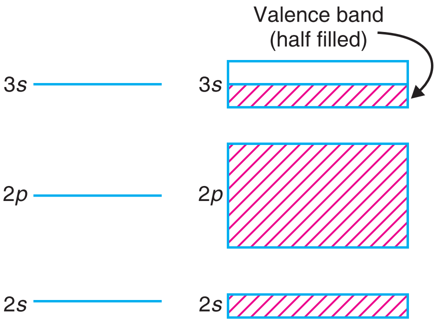


Fig. 4.10

(i) **Conductors.** Fig. 4.10 is a simplified diagram of the energy levels of a sodium atom ($1s^2\ 2s^2\ 2p^6\ 3s^1$) and the energy bands of solid sodium. A sodium atom has a single 3*s* electron in its valence shell. Thus if there are *N* atoms in a solid piece of sodium, its 3*s* valence band will contain *N* energy levels. The *N* valence electrons will fill in pairs only lower half of the energy levels of the band. Now suppose an electric field is applied across a piece of solid sodium. Then electrons in the valence band easily acquire additional energy to move to the higher unoccupied energy levels *within the same band* without crossing any energy gap. The additional energy is in the form of kinetic energy, and the moving electrons constitute an electric current. Sodium metal is, therefore, a good conductor of electricity. An unoccupied band into which electrons can pass is termed as 'conduction band'. In metals, the valence band itself is also the conduction band.

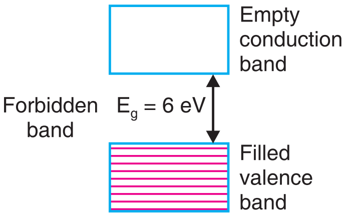


Fig. 4.11

(ii) **Insulators.** Fig. 4.11 is a diagram of the energy bands of diamond. The valence band is completely filled and the conduction band is empty at the absolute zero of temperature. There exists a large energy gap E_g of 6 eV between the valence and conduction bands. This means that atleast 6 eV of additional energy must be provided to an electron in a diamond crystal if it is to have any kinetic energy, since it cannot have an energy lying in the forbidden band. If an electric field is applied across diamond, the electrons in the valence band would not move within the band because there are no unoccupied levels in this band. They can, however, move to the higher empty band provided they get the required energy E_g to cross the forbidden band. An electric field cannot give this amount of energy to an electron in the solid. Thus no electron acquires additional kinetic energy in the electric field. Therefore, there will be no electric current and the solid will be an insulator.

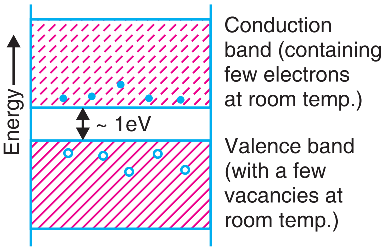


Fig. 4.12

(iii) **Semiconductors.** Silicon has a crystal structure resembling that of diamond. As in diamond, an energy gap separates the top of the filled valence energy band from a vacant higher band (conduction band). But the width of the forbidden band in silicon is only 1.1 eV (Fig. 4.12). At low temperatures, conductivity of silicon will be very low. Since the width of the energy gap is small, at ordinary room temperatures, a small proportion of electrons in the valence band have sufficient kinetic energy of thermal motion to cross the narrow energy gap and enter the conduction band above it. These electrons are sufficient to permit a limited amount of current to flow when an electric field is applied. Thus silicon has an electrical resistivity lying between that of conductors and insulators. Hence it is termed a semiconductor. Thus solids such as silicon and germanium are conductors at higher temperatures even though they are insulators at very low temperatures.

In a metallic conductor, increase in temperature impedes the flow of electrons due to the resulting greater thermal agitation of the lattice ions of the solid and hence the conductivity decreases. In semiconductors, this effect is masked by the ability of more and more electrons to overcome the energy gap as a result of their increased thermal energies at higher temperatures. The conductivity of a semiconductor, therefore, increases with temperature.

4.8 Optical Properties of Solids

The optical properties of solids and their energy band structures are closely connected. We can explain the following optical properties of solids on the basis of energy bands.

(*i*) All metals are opaque to light of all wavelengths. In a metal, the valence energy band is only partly filled. Therefore, the valence electrons can acquire additional energy, however small, to move to the unoccupied energy levels of the same band. Photons of visible light have energies between about 1 and 3 eV. Therefore, when light of any wavelength falls on a metal, the electrons absorb the light photons (energy) and are excited to the higher unoccupied energy levels. Hence metals are opaque to visible light.

(*ii*) Insulators are transparent to visible light. The energy gap in insulators is large. The electrons in an insulator need more than 3 eV of energy to jump across the forbidden band to the next allowed band. Insulators, therefore, cannot absorb photons of visible light. Hence insulators are transparent.

(*iii*) Semiconductors are transparent to infrared light but opaque to visible light. The energy gap in semiconductors are about the same in width as the energies of photons of visible light. The photons of visible light can excite the valence electrons to the upper empty conduction band. Hence visible light is absorbed in the semiconductor. Infrared light has lower frequencies and hence lower photon energies. So photons of infrared light cannot excite the valence electrons to the upper empty conduction band. Hence infrared light is not absorbed in the semiconductor.

4.9 Energy Bands: Alternative Analysis

We shall discuss qualitatively the motion of electrons in a periodic lattice and explain how it leads to the origin of the energy bands and the forbidden bands in solids.

In general, an electron in a solid moves in a region of *periodically-varying potential* (with the periodicity of the lattice) caused by the ion-cores situated at the lattice points, plus the average energy of all the other free electrons. Fig. 4.13 shows the potential energy (U) of an electron in a periodic array of positive ions. When the de Broglie wavelength of the electron ($= h/p$) corresponds to a periodicity in the spacing of the ions, the electron interacts strongly with the lattice and undergoes Bragg "reflection". This limits the electron to certain ranges of momenta and, correspondingly, to certain ranges of energy (energy bands).

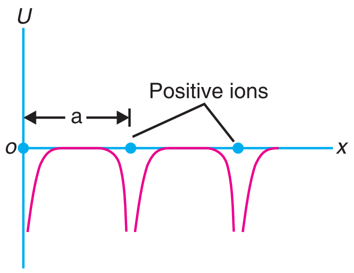


Fig. 4.13

The de Broglie wavelength of a free electron of momentum p is

$$\lambda = \frac{h}{p} \qquad \text{...(1)}$$

Let a be the spacing between the ions of the lattice. Unbound low-energy electrons can travel freely through a crystal since their wave lengths are long relative to the lattice spacing a.

More energetic electrons have wavelengths comparable with a. Such electrons are diffracted. The Bragg's condition for the reflection of waves is

$$2a \sin\theta = n\lambda, \quad n = 1, 2, 3, \ldots\ldots \qquad \ldots(2)$$

Here, the electron approaches the atomic plane at an angle θ.

The wave number k is

$$k = \frac{2\pi}{\lambda} = \frac{p}{\hbar} \qquad \ldots(3)$$

Bragg's formula in terms of k is

$$k = \frac{n\pi}{a\sin\theta}, \quad n = 1, 2, 3, \ldots\ldots \qquad \ldots(4)$$

Fig. 4.14 shows Bragg reflection in a two-dimensional square lattice. Bragg reflection from the vertical rows of ions occurs when $k_x = n\,\pi/a$. Similarly, reflection from the horizontal rows occurs when $k_y = n\,\pi/a$.

Brillouin Zones

The region in k-space (here an imaginary plane whose rectangular coordinates are k_x and k_y) that low-k electrons can occupy without being diffracted is called the *first Brillouin Zone*. The second Brillouin zone contains electrons with $k > \pi/a$ that do not fit into the first zone, yet which have sufficiently small wave numbers to avoid diffraction by the diagonal sets of atomic planes in Fig. 4.14. The second zone contains electrons with k values from π/a to $2\,\pi/a$ for electrons moving in the $\pm x$ and $\pm y$ directions. The first and second Brillouin zones of a two-dimensional square lattice are shown in Fig. 4.15. The range of k values between $-\pi/a$ and $+\pi/a$ defines the first Brillouin zone, those of k values between $-2\pi/a$ and $-\pi/a$ and between $+\pi/a$ and $+2\,\pi/a$ define the second Brillouin zone.

Brillouin Zones.

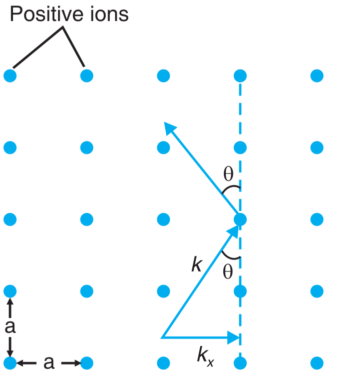


Fig. 4.14

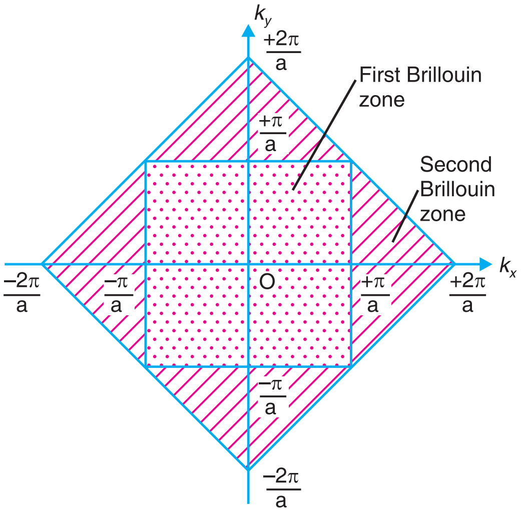


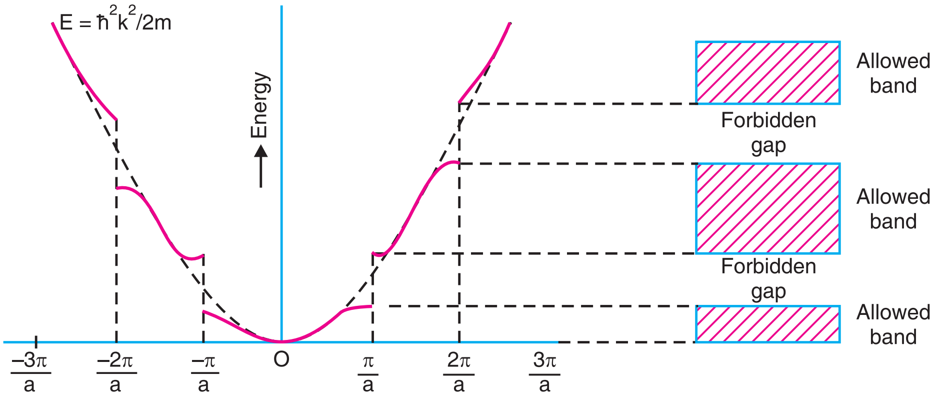
Fig. 4.15

Origin of Forbidden Bands

The energy of a free electron is related to its wave number k by

$$E = \frac{\hbar^2 k^2}{2m} \qquad \text{...(5)}$$

Fig. 4.16 shows how E varies with k in the x direction. The dashed curve is the free-electron parabola.



$E = \hbar^2k^2/2m$
Energy
Allowed band
Forbidden gap
Allowed band
Forbidden gap
Allowed band
$-\frac{3\pi}{a}$
$-\frac{2\pi}{a}$
$-\frac{\pi}{a}$
O
$\frac{\pi}{a}$
$\frac{2\pi}{a}$
$\frac{3\pi}{a}$

Fig. 4.16

Gaps in energy occur at values given by

$$k = \pm\frac{\pi}{a}, \pm\frac{2\pi}{a}, \pm\frac{3\pi}{a}, \ldots$$

At the above values of k we get energy gaps, whereas for values of k not near these values the energies are much like that of a free electron. At $k = \pi/a$, E has two values, the lower belonging to the first Brillouin zone and the higher to the second zone. There is a definite gap between the possible energies in the first and second Brillouin zones. This energy gap corresponds to a forbidden band.

The waves corresponding to values of k *not* satisfying the above condition travel almost freely. Those satisfying the condition, however, are reflected resulting in "standing" waves.

Let us consider electrons moving in the x direction. When $k = \pm\,\pi/a$, the waves are Bragg-reflected back and forth. So the only solutions of Schrodinger's equation consist of standing waves whose wavelength is equal to the periodicity of the lattice. There are two possibilities for these standing waves for $n = 1$:

$$\psi_1 = A \sin\frac{\pi x}{a}$$

$$\psi_2 = A \cos\frac{\pi x}{a}$$

Thus we get two *different* possible standing waves.

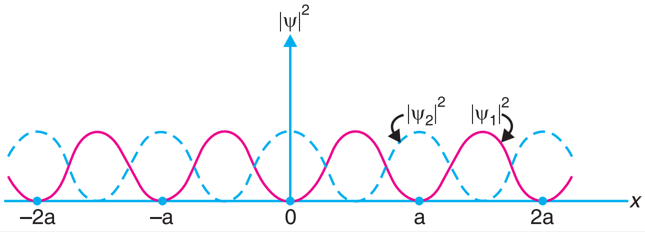


Fig. 4.17

The probability densities $|\psi_1|^2$ and $|\psi_2|^2$ for these waves are plotted in Fig. 4.17. Evidently $|\psi_1|^2$ has its minima at the lattice points occupied by the postive ions ($x = 0, \pm a, \pm 2a, ...$). $|\psi_2|^2$ has its maxima at the lattice points. The charge density corresponding to an electron wave function ψ is $e|\psi|^2$. The charge density in the case of ψ_1 is concentrated *between* the positive ions. In the case of ψ_2, it is concentrated *at* the positive ions. The potential energy of an electron in a lattice of positive ions is greatest midway *between* each pair of ions and least *at* the ions themselves. So an electron for $k = \pi/a$ can have two *different* energies E_1 and E_2 associated with the standing waves ψ_1 and ψ_2. No other solutions are possible when $k = \pm\ \pi/a$, and accordingly no electron can have an energy between E_1 and E_2. The difference $E_1 \sim E_2$ is the energy gap. This is the origin of all the energy gaps.

EXERCISE

1. Describe Millikan's method for the determination of electronic charge. Indicate the importance of this measurement.

2. Give an account of electronic theory of matter and explain how it accounts for the conduction of electricity through metals.

3. Derive an expression for the thermal conductivity from the free electron theory of metals.

4. Explain the principles on which the electron microscope works. Compare this instrument with an optical microscope.

5. Mention superiority of band theory of solids over free electron theory. Hence distinguish between conductors, semiconductors and insulators.

6. Discuss the motion of an electron in one-dimensional periodic potential and hence show that electron can not possess all possible energies but its energy spectrum contains alternate allowed and forbidden energy bands.

7. What voltage must be applied to an electron microscope to produce electrons of wavelength 0.40 Å?

[Andhra University, 2010]

Sol. de-Broglie wavelength λ is given by

$$\lambda = \frac{h}{p} = \frac{h}{\sqrt{(2mE)}} = \frac{h}{\sqrt{(2meV)}}$$

$E = eV$ where V is in volts

Here, $h = 6.6 \times 10^{-34}$ Js, $m = 9.1 \times 10^{-31}$ kg, $e = 1.6 \times 10^{-19}$ C

and $\lambda = 0.4 \times 10^{-10}$ m

$$\therefore \quad 0.4 \times 10^{-10} = \frac{6.6 \times 10^{-34}}{\sqrt{(2 \times 9.1 \times 10^{-31} \times 1.6 \times 10^{-19}\ V)}}$$

Solving, we get $V =$ **960 Volts.**

8. In germanium the energy gap is about 0.75 eV. Show that the crystal behaves as a transparent medium only for light of wavelength above 16533 Å. ($h = 6.6 \times 10^{-34}$ Js, $c = 3.0 \times 10^{8}$ ms^{-1}) **(Lucknow 85, 83)**

Sol. The energy of 16533-Å photons is

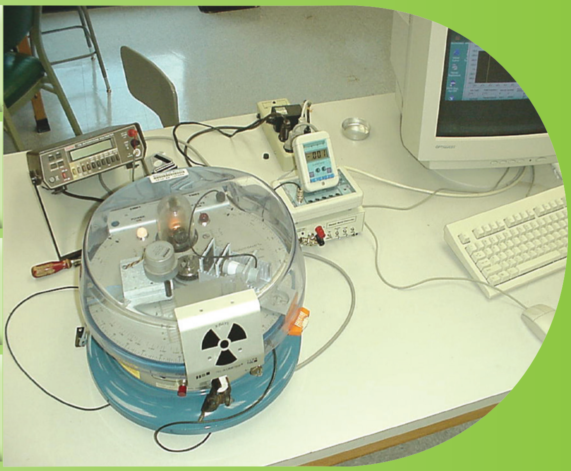
$$E = \frac{hc}{\lambda} = \frac{(6.6 \times 10^{-34}\ \text{Js})(3.0 \times 10^{8}\ \text{ms}^{-1})}{16533 \times 10^{-10}\ \text{m}} = 1.2 \times 10^{-19}\ \text{J} = 0.75\ \text{eV}.$$

The photon energy for light of wavelength above 16533 Å will be less than 0.75 eV which is the energy gap in germanium. Hence these photons will not be absorbed and the light will be transmitted through the crystal.

All photons for light below 16533 Å will have an energy more than the energy gap of 0.75 eV. So they will be absorbed in exciting the electrons from the valence band to the conduction band. Hence the crystal is opaque for light below 16533 Å.

CHAPTER

5



Positive Rays and Particle Properties of Waves

AT A GLANCE			
5.1	Discovery	**5.2**	Properties of Positive Rays
5.3	Positive Ray Analysis — Thomson's Parabola Method	**5.4**	Aston's Mass Spectrograph
5.5	Bainbridge's Mass Spectrograph	**5.6**	Dempster's Mass Spectrograph
5.7	Mass Defect and Packing Fraction	**5.8**	Polarisation of X-rays
5.9	Scattering of X-rays (Thomson's Formula)	**5.10**	Determination of the Number of Electrons per Atom
5.11	Failure of Classical Mechanics		

5.1 Discovery

If the cathode in a discharge tube is perforated, "*luminous rays*" are observed behind the cathode. These rays travel in a direction opposite to the cathode rays, *i.e.*, from the anode to the cathode. These rays produce fluorescence when they strike a fluorescent screen present at the back of the cathode (Fig. 5.1). They are deflected by electric and magnetic fields. The direction of deflection indicates that they are *positively charged*. These rays are called *positive rays* or *canal rays.*

These rays consist of positively charged ions produced by the ionisation of the gas atoms by the fast moving electrons of the cathode rays. These positive ions travel towards the cathode and emerge out through the perforations in the cathode. The positive ions are gas atoms which have lost one or more electrons. Hence the mass of a positive ion is almost the same as that of a neutral gas atom.

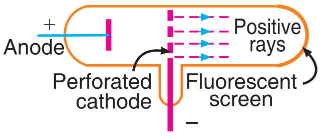


Fig. 5.1

5.2 Properties of Positive Rays

(1) These rays affect a photographic plate, produce fluorescence and penetrate thin aluminium foils.

(2) They are deflected by electric and magnetic fields. The direction of deflection indicates that they are positively charged. For the same electric or magnetic field, the magnitude of deflection suffered by positive rays is much smaller than that of cathode rays. Therefore, the positive ray particles are much heavier than the electrons.

(3) The velocities of all the positive ray particles is not the same. The velocities of these rays range from 10^5 to 10^6 ms^{-1}.

5.3 Positive Ray Analysis — Thomson's Parabola Method

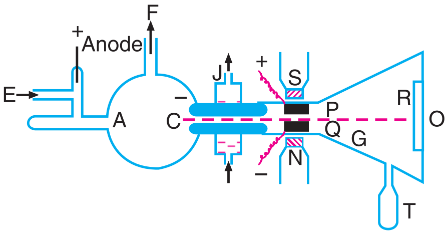


Fig. 5.2

Thomson determined the charge to mass ratio of positive ions by using the apparatus shown in Fig. 5.2. It consists of a discharge tube (*A*) in which the pressure of the gas is about 10^{-5} m of mercury. The anode is held in a side-tube. In order to ensure the supply of the gas under test, a steady stream of the gas is allowed to flow in through a capillary tube (*E*) and after circulating in *A* is pumped off at *F*. The cathode (*C*) is perforated with an extremely fine hole. The cathode is cooled by the water-jacket (*J*). The positive ions produced in *A* fly towards the cathode and those reaching it axially pass straight through the fine hole and emerge from the opposite end of the cathode as a narrow beam. This beam is then subjected to parallel electric and magnetic fields simultaneously. An electric field is applied between the plates *P* and *Q*. The electric field is perpendicular to the positive ray beam. *N* and *S* are the poles of a strong electromagnet. After passing through these fields, the beam enters a highly evacuated camera *G* and is received on a photographic plate *R*. A liquid air-trap (*T*) helps to keep the pressure in *G* quite low, even though pressure in *A* is comparatively large. The photographic plate, when developed, shows a series of parabolae.

Theory. Consider a positive ion of mass *M*, charge *E* and velocity *v*. When no electric or magnetic field is applied, the positive ion strikes the screen at *O*. This is called the *undeflected spot*.

Action of Electric Field. Let an electric field of strength *X* act over a length *l* of the path of the ion.

$$\left.\begin{array}{l}\text{Displacement of the ion in}\\ \text{passing through the electric field}\end{array}\right\} = s = \frac{1}{2}\left(\frac{XE}{M}\right)\left(\frac{l}{v}\right)^2.$$

After leaving the field, the ion moves in a straight line and finally strikes the plate at a distance *x* from *O*. *x* is proportional to *s* as well as to the distance between the field and the plate. Hence

$$x \propto \frac{XEl^2}{2Mv^2}$$

or
$$x = k_1 \frac{XE}{Mv^2} \qquad \text{...(1)}$$

Here k_1 is a constant.

Action of Magnetic Field. Suppose a magnetic field of strength *B* is applied over the same length *l* in the same direction as that of the electric field. The positive ion will now be deflected by this field in a direction at right angles to that in which it was deflected by the electric field. It will strike the plate at a distance *y* from *O* such that *oy* is perpendicular to *ox* in the plane of the plate.

$$\left.\begin{array}{l}\text{Displacement of the ion just}\\ \text{emerging from the magnetic field}\end{array}\right\} = s' = \frac{1}{2}\left(\frac{BEv}{M}\right)\left(\frac{l}{v}\right)^2 = \frac{BEl^2}{2Mv}$$

On emerging from the field, the ion moves in a straight line and finally strikes the plate at a distance *y* from *O*. *y* is proportional to *s'* as well as to the distance between the field and the plate.

Hence, $$y \propto \frac{BEl^2}{2Mv}$$

or $$y = k_2 \frac{BE}{Mv} \qquad \text{...(2)}$$

Here k_2 is another constant.

Action of Combined Electric and Magnetic Fields.

The combined effect of the two fields is found by eliminating v from (1) and (2).

Squaring (2) and dividing by (1),

$$\frac{y^2}{x} = \left(\frac{k_2^{\,2}}{k_1}\frac{B^2}{X}\right)\frac{E}{M} \qquad \text{...(3)}$$

B and X are constants. If E/M is constant, then by Eq. (3), $\frac{y^2}{x}$ = constant. This is the equation of a parabola. As Eq. (3) is independent of v, particles of same E/M but of different velocities will fall on different points on the *same parabola.*

$\frac{y}{x} = \frac{k_2}{k_1}\frac{B}{X} v$. *i.e.*, $\frac{y}{x} \propto v$. Thus, the position of any individual particle on the parabola will depend on the velocity of the particle.

The ions having different values of E/M will lie along the different parabolas.

For one direction of the magnetic field, one half of the parabola is traced. Reversing the magnetic field, the other half is also traced. When the full parabola is traced, it is easy to draw the axis of symmetry (X-axis).

Determination of *E/M*. The value of E/M can be calculated from Eq. (3) by measuring the coordinates x and y for a point on the parabola, evaluating the constants k_1 and k_2 for the apparatus and knowing B and X.

Determination of Mass. The mass of a positive ion is determined in terms of the mass M_1 of the standard hydrogen ion. A small trace of hydrogen is always present in all samples of gases. Hydrogen, being the lightest element, gives the outermost parabola. Let I and II represent the parabolic traces due to ions of the gas of mass M_2 and the hydrogen ions of mass M_1 respectively (Fig. 5.3). An ordinate a,b,c,d,e is drawn. It cuts the two parabolas at a, b, d, e and the X-axis at c. Let ac and bc represent the two values of y corresponding to a constant value of x on these parabolas. Let us assume that both ions have the same charge. Then from Eq. (3),

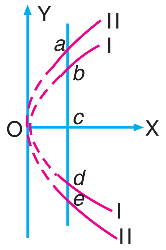


Fig. 5.3

$$\frac{ac^2}{x} = \left(\frac{k_2^{\,2}}{k_1}\frac{B^2}{X}\right)\frac{E}{M_1}$$

and $$\frac{bc^2}{x} = \left(\frac{k_2^{\,2}}{k_1}\frac{B^2}{X}\right)\frac{E}{M_2}$$

or $$\frac{M_2}{M_1} = \frac{ac^2}{bc^2} = \left(\frac{ae}{bd}\right)^2$$

The lengths ae and bd can be measured on the photograph. Hence the parabolic traces enable us to compare the masses of different ions with hydrogen used as a standard.

$$M_2 = \left(\frac{ae}{bd}\right)^2 M_1.$$

Discovery of Stable Isotopes. Using neon gas in his apparatus, Thomson obtained two parabolas for the gas itself, a strong one corresponding to a mass 20 and a much weaker one corresponding to a mass 22. The intensity ratio of the two traces was 9 : 1 which gave the relative abundance of the two isotopes. Thomson, therefore, suggested that neon could exist in the form of two *isotopes*, chemically indistinguishable but with different masses 20 and 22. The actual observed atomic weight (20.2) of neon is the weighted mean of the masses of these two isotopes.

This established the existence of stable isotopes.

Limitations of the Parabola Method:

1. Due to the velocity dispersion, each parabolic trace is of very low intensity.
2. The traces on the photographic plate are blurred and have no definite edges. Hence accurate measurements are not possible.
3. The influence of secondary rays makes analysis difficult.

5.4 Aston's Mass Spectrograph

Aston's Mass Spectrograph.

The apparatus used by Aston is shown in Fig. 5.4. The stream of positive ions obtained from a discharge tube is rendered into a fine beam by passing it between two narrow slits S_1 and S_2. This beam enters the electric field between the metal plates P_1 and P_2. Due to the action of the electric field (X), all positive ions having the same value of E/M are not only deviated by an angle θ from the original path but are dispersed by an angle $d\theta$ due to their different velocities. The beam is then allowed to pass through a magnetic field M acting at right angles to the electric field so that it produces a deflection of the beam in the same plane. The magnetic field deviates the particles by an angle ϕ and reconverges them by $d\phi$. The direction and magnitude of the field is so adjusted that it produces a deviation of the beam in the opposite direction and brings all ions having the same value of E/M, eventhough differing in velocities, to a focus at one point F. Ions having different values of E/M are brought to focus at different points on the photographic plate. The condition required for such a focusing may be derived as follows:

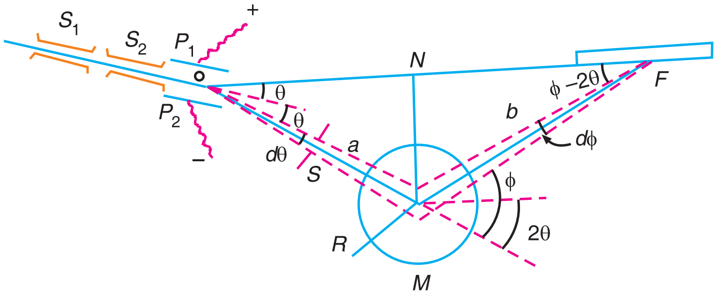


Fig. 5.4

Theory : To obtain the condition of focusing. Consider a group of ions having the same value of E/M, but moving with different velocities. Let θ and ϕ be the mean angles of deviation of the group of ions in the electric and magnetic fields respectively. Let $d\theta$ be the dispersion angle due to the electric field and $d\phi$ the convergence angle due to the magnetic field.

(*i*) Let X = The strength of the electric field,

E = Charge of the ion.

M = Mass of the ion and

v = Velocity of the ion.

l_1 = the length of the electric field

and d_1 = The linear displacement of an ion from its path due to the electric field.

then
$$d_1 = \frac{1}{2}\frac{XE}{M}\frac{l_1^2}{v^2}$$

$$\left.\begin{array}{r}\text{The angular deviation}\\ \text{of the beam}\end{array}\right\} = \theta = \frac{d_1}{l_1} = \frac{1}{2}\frac{XE}{Mv^2}l_1 \qquad ...(1)$$

$$\left.\begin{array}{r}\text{Angle of dispersion}\\ \text{of the beam}\end{array}\right\} = d\theta = \frac{-XEl_1}{M}\frac{dv}{v^3} \qquad ...(2)$$

$$\therefore \quad \frac{d\theta}{\theta} = -\frac{2dv}{v} \qquad ...(3)$$

(*ii*) Similarly, if d_2 = displacement of an ion from its path due to the magnetic field of strength B, and l_2 = the length of path of the ion in the magnetic field,

$$d_2 = \frac{1}{2}\frac{BEv}{M}\frac{l_2^2}{v^2}$$

$$\left.\begin{array}{r}\text{The deviation produced by the magnetic}\\ \text{field in the opposite direction}\end{array}\right\} = \phi = \frac{d_2}{l_2} = \frac{1}{2}\frac{BE}{Mv}l_2 \qquad ...(4)$$

$$\therefore \quad d\phi = -\frac{1}{2}\frac{BEl_2}{M}\frac{dv}{v^2} \qquad ...(5)$$

Hence
$$\frac{d\phi}{\phi} = -\frac{dv}{v} \qquad ...(6)$$

Comparing (3) and (6),
$$\frac{d\theta}{\theta} = \frac{2d\phi}{\phi}$$

or
$$\frac{d\phi}{d\theta} = \frac{\phi}{2\theta} \qquad ...(7)$$

Let $a = OR$ = The distance between the two fields. The width of the selected group of ions at $R = a\ d\theta$. If there had been no magnetic field, the width of the group after travelling a further distance b would be $(a + b)\,d\theta$. The magnetic field produces a convergence $d\phi$ and brings the group of ions to a focus at a distance $RF = b$.

∴ The condition of focusing is given by

$$(a + b)\,d\theta = b\,d\phi \text{ or } \frac{a+b}{b} = \frac{d\phi}{d\theta} = \frac{\phi}{2\theta} \text{ [From Eq. 7]}$$

or
$$\frac{b}{a} = \frac{2\theta}{\phi - 2\theta} \qquad ...(8)$$

This expression represents a straight line drawn from O, making an angle 2θ with the direction of the beam deviated by the electric field. If $\phi = 2\theta$, the rays do not converge (since $b = \infty$). However, if $\phi = 4\theta$, $b = a$ and the convergent beam can be easily photographed. The ions are focused at F such that $OR = RF$.

Detection of isotopes. To determine the masses of isotopes of an element, the element is mixed with a number of other elements whose atomic masses are known accurately. The traces of all of them are obtained on the same photographic plate. The distances of the traces of the known masses are measured from a given reference point on the plate. A calibration graph is drawn connecting the distances with the atomic masses. By measuring the distances of the traces of the isotopes of the element from the same reference point, their atomic masses are obtained from the calibration graph.

Advantages : 1. All the particles having the same value of E/M are focussed at a single point unlike Thomson's method in which they are spread out into a parabola. Hence, the intensities of the lines on the photographic plate in Aston's mass spectrograph are large, while the parabolic traces on the photographic plate in Thomson's arrangement are feeble.

2. The intensity of a line in the mass spectrum is proportional to the total number of particles of that mass. Hence a rough idea of the relative abundance of various isotopes of an element can be made by this method.

Limitations : 1. The mass scale is not linear.

2. Owing to the polarisation of the electrodes of the electric field, the traces are not quite straight but slightly curved.

Example. *The distance between traces corresponding to masses 12 and 16 in an Aston's mass spectrograph is 4.8 cm. Calculate the mass of the particle whose trace is at a distance of 8.4 cm from the trace of mass 16.*

Sol. Let x be the distance of the fiduciary mark from the trace of mass 16.

$$\therefore \qquad x \propto 16. \text{ Also, } x - 4.8 \propto 12.$$

$$\text{Hence} \qquad \frac{x}{x-4.8} = \frac{16}{12} \quad \text{or} \quad x = 19.2 \text{ cm.}$$

If M be the mass of the particle whose trace is at a distance of 8.4 cm from the trace of mass 16, we have

$$19.2 \pm 8.4 \propto M \text{ and } 19.2 \propto 16.$$

$$\frac{M}{16} = \frac{19.2 \pm 8.4}{19.2} = 1 \pm \frac{8.4}{19.2} = 1 \pm \frac{7}{16}$$

$$\text{Hence} \qquad M = 16 \pm 7 = 23 \text{ or } 9.$$

5.5 Bainbridge's Mass Spectrograph

The apparatus is shown diagrammatically in Fig. 5.5. The beam of positive ions produced in a discharge tube is collimated by two slits S_1 and S_2 and enter a "*velocity selector*". The velocity selector consists of (1) a steady electric field X maintained at right angles to the ion beam between two plane parallel plates P_1 and P_2 and (2) a magnetic field B. The magnetic field is produced by an electromagnet represented by the dotted circle. The magnetic field B is perpendicular to X and the ion beam. The electric field and magnetic field of the velocity selector are so adjusted that the deflection produced by one is nullified by the deflection produced by the other. If X and B are the electric intensity and magnetic induction, then $Xe = Bev$ or $v = X/B$.

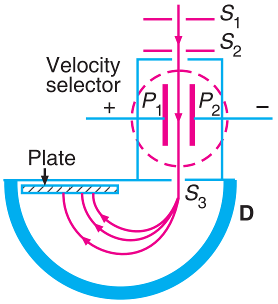


Fig. 5.5

Only those ions, having this velocity v, alone pass through the entry slit S_3 to enter the evacuated chamber D. Thus all ions entering D must have the same velocity. The positive ions which enter into D are subjected to a strong uniform magnetic field of intensity B', perpendicular to its path. The force acting on each ion will be $B'ev$.

Ions with different masses trace circular paths of different radii given by

$$R = \frac{Mv}{B'e} \left(\because\ B'ev = \frac{Mv^2}{R} \right)$$

$$\therefore \qquad \frac{e}{M} = \frac{v}{B'R} = \frac{X}{BB'R} \left(\because\ v = \frac{X}{B} \right).$$

Since v and B' are constant quantities, $\frac{e}{M} \propto \frac{1}{R}$.

After describing semicircles, the ions strike a photographic plate.

Now, $M = \frac{B'eR}{v}$. If e is the same for all ions, then $M \propto R$.

So we get a linear mass scale on the photographic plate. It will be seen that ions of different masses strike the photographic plate at different points, thus giving a typical mass spectrum.

Advantages : (1) Since a linear mass scale is obtained, accuracy of measurements is increased.

(2) The sensitivity depends on the strength of the deflecting magnetic field B' and the field area of the chamber D. Bainbridge used a magnetic field of 1.5 weber/m^2 over a semicircle of radius 0.2m. He found a definite increase in resolving power over Aston's apparatus. The ten isotopes of tin were resolved by this instrument.

EXAMPLE 1. *A beam of positively charged particles go through a velocity selector where an electric field of* 6×10^5 *volt/metre and a magnetic induction of* 2×10^{-2} *weber/m*^{2} *are applied perpendicular to each other. Then the particles go through another magnetic induction of* 1.25 *weber/m*^{2} *applied perpendicular to the path. If the diameter of the path taken is one metre, calculate the specific charge of the particles.*

SOL. We have, $\frac{e}{M} = \frac{X}{BB'R}$. Here, $X = 6 \times 10^5$ volt/metre;

$B = 2 \times 10^{-2}$ weber/m^2, $B' = 1.25$ weber/m^2 and $R = 0.5$ m.

$$\therefore \qquad \frac{e}{M} = \frac{6 \times 10^5}{(2 \times 10^{-2})\ (1.25)\ 0.5} = 4.8 \times 10^7 \text{ coulomb/kg.}$$

EXAMPLE 2. *Singly-ionised Mg atoms enter a Bainbridge mass spectrograph with the velocity selector having electric and magnetic fields respectively of* 30 *kilovolts/metre and* 0.1 *tesla. Calculate the radii of the path followed by the three isotopes of masses* 24, 25 *and* 26 *when the deflecting magnetic field is* 0.5 *tesla.*

SOL. The mass of a nucleon is 1.67×10^{-27} kg.

The ions emerging from the velocity selector have a velocity given by

$$v = \frac{X}{B} = \frac{30000}{0.1} = 3 \times 10^5 \text{ ms}^{-1}$$

Now, if B' is the "deflecting" magnetic field, the radius of the path of an ion of mass M, velocity v and charge e is given by

$$R = \frac{Mv}{B'e} = \frac{3 \times 10^5\ M}{(0.5)\ (1.6 \times 10^{-19})} = 3.75 \times 10^{24}\ M.$$

Let R_1, R_2 and R_3 be the radii of isotopes of mass numbers 24, 25 and 26 respectively.

$$\text{Mass of Mg}^{24}\text{ nucleus} = 24 \times (1.67 \times 10^{-27})\text{ kg.}$$

$$\therefore \quad R_1 = (3.75 \times 10^{24}) \times (24 \times 1.67 \times 10^{-27}) = 0.150\text{ m.}$$

$$R_2 = (3.75 \times 10^{24}) \times (25 \times 1.67 \times 10^{-27}) = 0.156\text{ m.}$$

$$R_3 = (3.75 \times 10^{24}) \times (26 \times 1.67 \times 10^{-27}) = 0.163\text{ m.}$$

Example 3. *In a Bainbridge mass spectrograph, singly ionised atoms of* Ne^{20} *pass into the deflection chamber with a velocity of* 10^5 ms^{-1}. *If they are deflected by a magnetic field of flux density* 0.08 *T, calculate the radius of their path and where* Ne^{22} *ions would fall if they had the same initial velocity.*

Sol. For an ion of Ne^{20} moving in a magnetic field B we can write

$$Bev = \frac{Mv^2}{R_{20}} \text{ or } R_{20} = \frac{Mv}{Be}$$

Here, $M = 20 / (6.023 \times 10^{26})$ kg; $v = 10^5\text{ ms}^{-1}$, $B = 0.08\ T$ and $e = 1.602 \times 10^{-19}$ C.

$$\therefore \quad R_{20} = \frac{(20/6 \times 10^{26}) \times 10^5}{0.08 \times (1.602 \times 10^{-19})} = 0.26\text{ m}$$

If e is the same for all ions, then $R \propto M$. The radius of the path followed by Ne^{22} is given by

$$\frac{R_{22}}{R_{20}} = \frac{22}{20} \text{ or } R_{22} = \frac{22}{20} \times R_{20} = \frac{22}{20} \times 0.26 = 0.286\text{ m.}$$

$$2R_{22} - 2R_{20} = 2 \times 0.026\text{ m} = 0.052\text{ m.}$$

Hence Ne^{22} ions would fall 5.2 cm away from the Ne^{20} ions.

5.6 Dempster's Mass Spectrograph

The experimental arrangement is shown in Fig. 5.6. The anode is a metal cylinder A with its front surface C coated with a salt of the element under test and heated electrically. A filament, if electrically heated by the battery B_1, emits electrons. By maintaining the filament at a P.D. of about 50 volts with respect to A by using another battery B_2, the electrons are made to bombard the heated salt with the result that the anode emits positively charged ions of the element. These ions are collimated into a narrow beam by the slit S_1. Then the positive ions are accelerated towards the slit S_2 by a variable P.D. V maintained between S_1 and S_2.

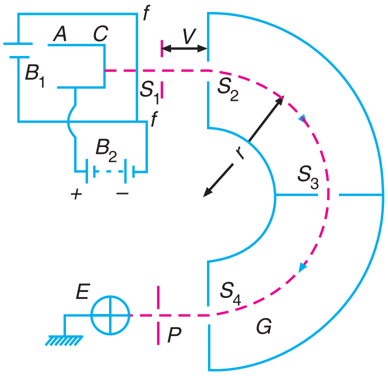


Fig. 5.6

We know that when ions of mass M and charge e are accelerated through a P.D. V, they acquire a velocity v given by

$$\frac{1}{2}Mv^2 = eV$$

or
$$v = \sqrt{\frac{2eV}{M}} \qquad \text{...(1)}$$

The velocity of the ions v is very large compared with their initial speeds of emission. The ions then enter a vacuum chamber G. On entering the space G, the ions are subjected to a magnetic field of flux density B directed at right angles to the plane of the diagram. The ions are deflected through a semicircle towards the slit S_4 defined by the three slits S_2, S_3 and S_4. The magnetic field of flux density B makes an ion of mass M entering with the velocity v traverse a circular path of radius r given by

$$Bev = \frac{Mv^2}{r}$$

or $$r = \frac{Mv}{Be} \qquad \text{...(2)}$$

But, from (1), $v = (2eV/M)^{1/2}$

$$\therefore \qquad r = \frac{Mv}{Be} = \frac{M}{Be}\left(\frac{2eV}{M}\right)^{\frac{1}{2}} = \left(\frac{2MV}{B^2 e}\right)^{\frac{1}{2}}$$

or $$r^2 = \frac{2MV}{B^2 e}$$

$$\therefore \qquad \frac{e}{M} = \frac{2V}{B^2 r^2} \qquad \text{...(3)}$$

For this particular value of V, only ions with the specific charge given by equation (3) are collected by the electrode P and the current is registered by the electrometer E.

From the equation $r = \left(\frac{2MV}{B^2 e}\right)^{\frac{1}{2}}$, it is clear that if $\left(\frac{M}{e}\right)\frac{V}{B^2}$ is kept constant, the radius r of the path of these ions will remain unchanged. Hence ions of different e/M can be made to traverse a circular path of fixed radius by adjusting V and B.

Determination of the masses of the isotopes: In the actual experiment, the magnetic field (B) is kept constant and the ionic current is measured by the electrometer for various values of V. A graph is plotted with this current against V. The graph will show sharp peaks at those values of V corresponding to the value of M given by Eq. (3). Fig. 5.7 gives the curves for potassium, showing the two isotopes of mass numbers 39 and 41. The relative abundance of the isotopes can be inferred from the relative heights of the peaks.

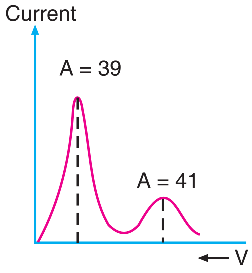


Fig. 5.7

Example. *In a Dempster mass spectrograph, a 1000 volt accelerating potential brings* $_{12}Mg^{24}$ *on the slit. What is the potential required to bring* $_{12}Mg^{25}$ *on the slit, the magnetic field being kept constant?*

Sol. The basic equations are: $eV = \frac{1}{2}Mv^2$ and $Bev = Mv^2/r$.

Thus if V is the accelerating potential required to bring an ion of mass M on the slit, we have

$$V = \frac{B^2 r^2 e}{2M} = \frac{k}{M}, \quad \text{where } k \text{ is a constant.}$$

For $M = 24$, $V = 1000$ volts. Thus $1000 = \frac{k}{24}$ or $k = 24000$.

For $M = 25$, we shall have $V = \frac{k}{M} = \frac{24000}{25} = 960$ volts.

Uses of Mass Spectrographs: 1. Mass spectrographs have been used by many research workers to determine the isotopic masses of almost all elements on the periodic table. Their results show that there are more than 300 different isotopes occurring in nature.

2. Mass spectrographs have been employed to separate small quantities of pure isotopes. Large equipments have been built to separate large quantities of uranium –235 which is used in the manufacture of atomic bombs and production of nuclear power.

5.7 Mass Defect and Packing Fraction

"Mass defect" of a given isotope is defined as the difference between the experimentally measured mass of the isotope (M) *and its mass number* (A).

Mass defect = $\Delta = M - A$.

The packing fraction of the isotope is defined as the ratio of its mass defect ($M - A$) *to its mass number* (A).

Packing fraction = $(M - A)/A$.

Since atomic masses are measure relative to C-12, the packing fraction for this isotope is zero.

Binding Energy

Definition. *When Z protons and N neutrons combine to make a nucleus, some of the mass* (Δm) *disappears because it is converted into an amount of energy* $\Delta E = (\Delta m)\, c^2$. *This is called the binding energy* (*B.E.*) *of the nucleus.*

$$\text{B.E.} = (Zm_p + Nm_n)\, c^2 - M_n c^2$$

where m_p = mass of the proton, m_n = mass of the neutron and M_n = nuclear mass.

Atomic Masses. Atomic masses refer to the masses of neutral atoms, not of bare nuclei. Thus an atomic mass always includes the masses of its Z electrons. Atomic masses are expressed in *mass units* (u), which are so defined that the mass of a ${}^{12}_{6}\text{C}$ atom is exactly $12u$.

The value of a mass unit is

$$1u = 1.66054 \times 10^{-27}\ \text{kg}.$$

The energy equivalent of a mass unit is 931.49 MeV.

Definition. *The atomic mass unit (u) is defined as one-twelfth of the mass of the neutral carbon C-12 atom.*

5.8 Polarisation of X-rays

Fig. 5.8 shows the experimental arrangement for showing the polarisation of X-rays. A beam of X-rays travelling along AZ is allowed to strike at a scattering sheet Z. This scatters X-rays in all directions. With the help of slits S_3 and S_4, a beam in the direction ZO which is perpendicular to the direction OX is isolated. The beam ZO strikes another scatterer at O. An ionization chamber $I.C.$ is placed to receive X-rays in the directions OX and OY. It is observed that the intensity of X-rays scattered in the direction OX is maximum and is nearly zero in the OY direction. The beam ZO is said to be polarized. Thus X-rays from the X-ray tube become completely polarized after having been scattered in the direction at right angles to the initial direction. The scatterer at Z behaves like a polarizer and that at O as an analyser.

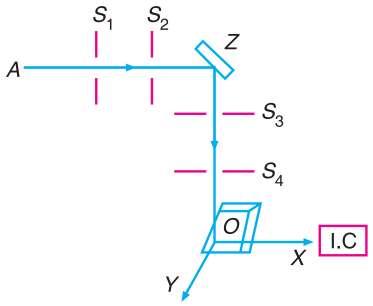


Fig. 5.8

5.9 Scattering of X-rays (Thomson's Formula)

When an electromagnetic wave interacts with a single free electron, the electron performs simple harmonic motion and in so doing reradiates. The radiations thus emitted by the vibrating electron under the influence of the incident X-ray beam are called *scattered X-rays*.

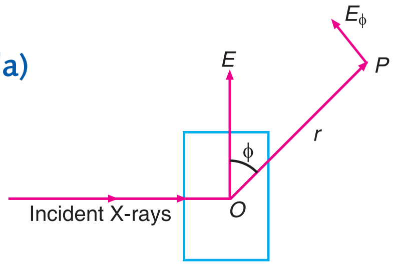


Fig. 5.9

At any point O within the material let E and B be the electric and magnetic vectors of the incident X-ray beam (Fig. 5.9). Both are perpendicular to the incident direction of the beam. The electron at the point O will experience an acceleration (a) in the direction of the electric field E.

$$a = eE/m \qquad \text{... (1)}$$

According to classical electromagnetic theory, an accelerated charge radiates electromagnetic waves. The electric vector of the scattered X-rays at a point P with polar coordinates r and ϕ is given by

$$E_\phi = \frac{ea\sin\phi}{4\pi\epsilon_0 c^2 r} = \frac{Ee^2\sin\phi}{4\pi\epsilon_0 mc^2 r} \qquad \text{...(2)}$$

$$\therefore \qquad \text{——} = \frac{e^2\sin\phi}{4\pi\epsilon_0 mc^2 r} \qquad \text{...(3)}$$

Intensity is proportional to the square of the electric vector.

So the ratio of the scattered intensity at P to the incident intensity is

$$\frac{I_\phi}{I} = \left(\frac{E_\phi}{E}\right)^2 = \frac{e^4\sin^2\phi}{(4\pi\epsilon_0)^2 m^2 c^4 r^2} \qquad \text{...(4)}$$

Consider a case where the incident beam is along the X-axis and its electric vector E is in the YZ plane (Fig. 5.10). Resolve E into two perpendicular components along the Y and Z axes.

$$E^2 = E_y^2 + E_z^2$$

$$\therefore \qquad I = I_y + I_z$$

Here I_y and I_z are the components of the wave along the Y and Z axes.

The incident beam is unpolarized.

$$\therefore \qquad I_y = I_z = \frac{I}{2} \qquad \text{...(5)}$$

Let I_1 and I_2 be the intensities of the scattered beam at the point P due to the y and z components.

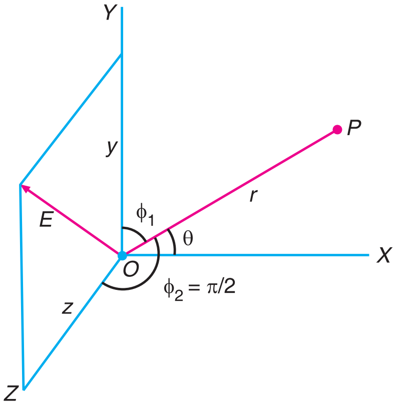


Fig. 5.10

$$\therefore \qquad I_1 = I_y\left[\frac{e^4\sin^2\phi_1}{(4\pi\epsilon_0)^2 m^2 c^4 r^2}\right]$$

$$I_y = \frac{I}{2} \text{ and } \sin^2\phi_1 = \cos^2\theta$$

Here θ is the angle between r and the x-axis.

$$\therefore \qquad I_1 = \frac{I}{2}\left[\frac{e^4\cos^2\theta}{(4\pi\epsilon_0)^2 m^2 c^4 r^2}\right] \qquad \text{... (6)}$$

Similarly

$$I_2 = \frac{I}{2}\left[\frac{e^4\sin^2\phi_2}{(4\pi\epsilon_0)^2 m^2 c^4 r^2}\right]$$

But

$$I_z = \frac{I}{2} \text{ and } \sin^2\phi_2 = 1.$$

Here the angle ϕ_2 between OP and OZ is always $\pi/2$.

$$\therefore \qquad I_2 = \left(\frac{I}{2}\right)\left[\frac{e^4}{(4\pi\epsilon_0)^2 m^2 c^4 r^2}\right] \qquad \text{... (7)}$$

The total intensity of the scattered wave at P is given by

$$I_e = I_1 + I_2$$

$$I_e = \left(\frac{I}{2}\right)\left[\frac{e^4[1+\cos^2\theta]}{(4\pi\epsilon_0)^2 m^2 c^4 r^2}\right] \qquad \text{... (8)}$$

Let n = number of electrons per unit volume of the scattering medium.

We assume that each electron is equally effective in producing scattered X-rays, irrespective of the presence of the others.

The intensity of the scattered X-rays at the point P, per unit volume of the scatterer is

$$I_s = n\, I_e$$

$$I_s = \frac{I}{2}\left[\frac{ne^4\,(1+\cos^2\theta)}{(4\pi\epsilon_0)^2\, m^2\, c^4\, r^2}\right] \qquad ...(9)$$

Equation (9) represents Thomson's scattering formula.

5.10 Determination of the Number of Electrons per Atom

Consider a sphere of radius r with centre O (Fig. 5.11). The scatterer is at O.

Take two points A and B on the surface. The element AB subtends an angle $d\theta$ at O.

Surface area of the element

$$= 2\pi\,(AC)\,(AB)$$

$$= 2\pi\,(r\sin\theta)\, r\, d\theta$$

$$= 2\pi\, r^2 \sin\theta\, d\theta$$

The total amount of energy scattered per unit volume of the scatterer in unit time is

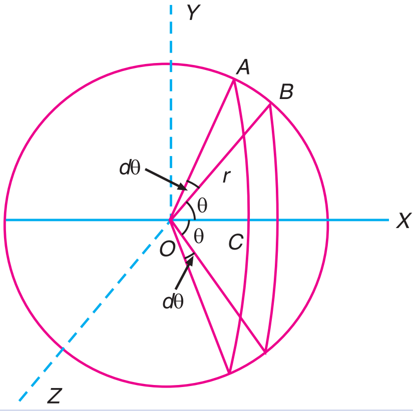


Fig. 5.11

$$U_s = \int_0^{\pi} I_s\,(2\pi r^2 \sin\theta)\, d\theta$$

$$U_s = \left(\frac{I}{2}\right)\left[\frac{2\pi n e^4 r^2}{(4\pi\epsilon_0)^2\, m^2\, c^4\, r^2}\right]\int_0^{\pi}(1+\cos^2\theta)\,\sin\theta\, d\theta \qquad ...(1)$$

$$U_s = \left(\frac{8\pi}{3}\right)\left[\frac{I\, n e^4}{(4\pi\epsilon_0)^2\, m^2\, c^4}\right] \qquad ...(2)$$

The linear scattering coefficient σ *of the material is*

$$\sigma = \frac{U_s}{I} = \frac{8\pi}{3}\left[\frac{n e^4}{(4\pi\epsilon_0)^2\, m^2\, c^4}\right] \qquad ...(3)$$

$$\therefore \qquad n = \left[\frac{3}{8\pi}\right]\left[\frac{(4\pi\epsilon_0)^2\, m^2\, c^4}{e^4}\right]\sigma \qquad ...(4)$$

By measuring σ, *the value of* n can be determined.

n / ρ represents the number of electrons per unit mass.

$$\therefore \qquad \frac{n}{\rho} = \left[\frac{3}{8\pi}\right]\left[\frac{(4\pi\epsilon_0)^2\, m^2\, c^4}{e^4}\right]\left(\frac{\sigma}{\rho}\right) \qquad ...(5)$$

The ratio of σ to the density ρ of the scatterer is the *mass scattering coefficient* σ_m.

Barkla and his coworkers found that the value of σ_m is 0.02 m^2/kg for carbon.

From Eq. (5), the number of electrons per kilogram is

$$\frac{n}{\rho} = 3 \times 10^{26} \text{ electrons/kg}$$

The number of carbon atoms per kilogram is

$$\frac{N_A}{A} = \frac{6.02 \times 10^{26} \text{ atoms/k mole}}{12 \text{ kg/k mole}} = 5 \times 10^{25} \text{ atoms/kg}$$

$$\left.\begin{array}{l}\text{The number of electrons} \\ \text{per carbon atom}\end{array}\right\} = \frac{3 \times 10^{26}}{5 \times 10^{25}} = 6 \cdot$$

Note

The ratio of the power scattered to the primary intensity is called the *scattering cross section* (or *scattering coefficient*) of the free electron, designated by σ_e.

$$\sigma_e = \left(\frac{8\pi}{3}\right)\left(\frac{e^2}{4\pi\epsilon_0 mc^2}\right)^2$$

5.11 Failure of Classical Mechanics

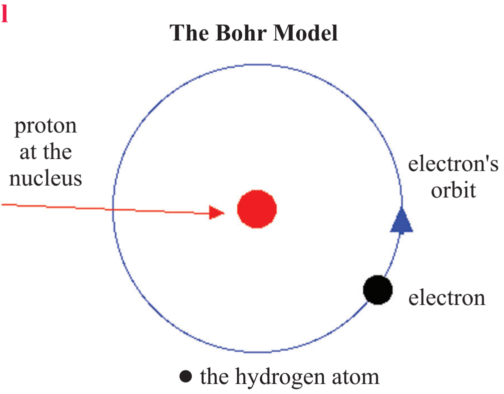
Classical mechanics failed to explain phenomena like atomic spectra, blackbody radiation, photoelectric effect, Compton effect and specific heat capacity of solids. The inadequacy of classical mechanics led to the development of *Quantum Mechanics*.

(*i*) The Hydrogen Atom and the Bohr Model

As the first example of the failure of classical physics to account for observed phenomena, we consider the case of the hydrogen atom. Rutherford model failed to explain two main observational features of the hydrogen atom:

(*a*) its stability and (*b*) the spectrum of its radiation. Let us consider these one at a time.

(*a*) An electron in a curved orbit is accelerated and hence must radiate. As it radiates its energy away, the radius of its orbit must decrease until eventually it collapses into the nucleus. Thus, the atom cannot be stable. But most of the atoms are stable.



Bohr's model of the hydrogen atom.

(*b*) The second discrepancy involves the observed radiation spectrum. The frequency of the radiated energy should be the same as the orbiting frequency. As the electron orbit collapses, its orbiting frequency increases continuously. We might thus expect the spectrum of radiation emitted by excited hydrogen atoms to be continuous. In contrast, the experimentally observed spectrum consists of families of discrete lines.

Bohr provided an explanation for both the spectral discreteness and the observed stability. He proposed that in solving for the orbital motion of the electron in its hydrogenic orbit one should impose an added condition:

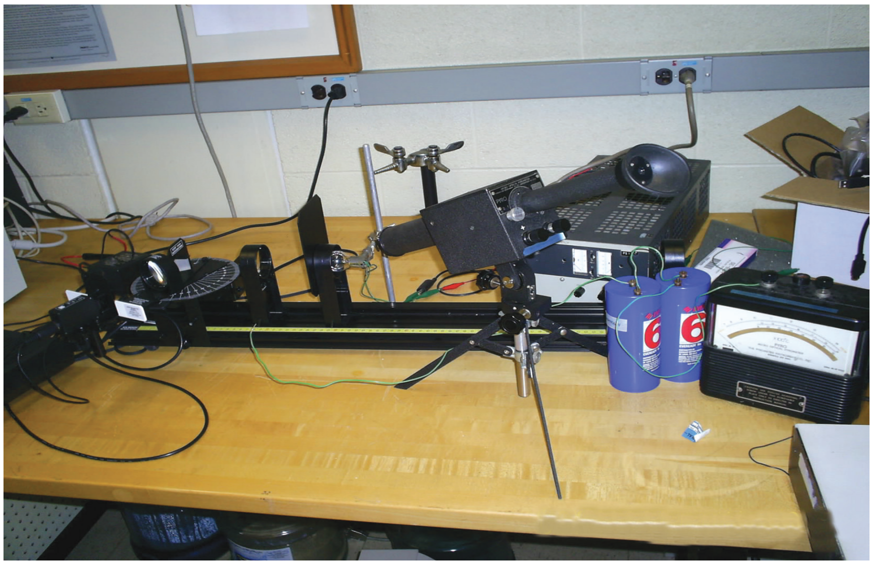
The angular momentum of the electron must be equal to some integer multiple of $\hbar$.

$$l = n\hbar \ (n = 1, 2, 3...)$$

The quantisation implies that the laws of classical mechanics and of classical electro-magnetism are inapplicable at the atomic level.

(*ii*) Black Body Radiation

The observed variation of the spectral intensity $I(\nu)$ (power per unit area per unit frequency) of blackbody radiation as a function of frequency ν is shown in Fig. 5.12. From the curves we note that



Black Body Radiation.

(*i*) The intensity reaches a maximum at some frequency ν_m.

(*ii*) The frequency ν_m, as well as the height of the peak, increase with temperature.

The application of statistical thermodynamics and the ordinary laws of mechanics and electromagnetic theory led to the Rayleigh-Jeans formula

$$I(v) = \frac{2\pi}{c^2} kT\nu^2 .$$

This law, except for very low frequencies, is in total disagreement with experimental results. The law predicts an infinite amount of radiated intensity. Actually, the total radiated intensity is finite.

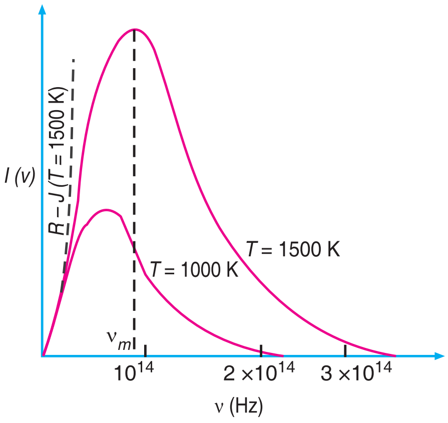


Fig. 5.12

Max Planck resolved this controversy by postulating that the exchange of energy between atoms and radiation involves *discrete* amounts of energy. At a given frequency ν, the smallest amount of energy that can be exchanged is equal to

$$E = h\nu .$$

Only multiples of $h\nu$ are involved in the interaction.

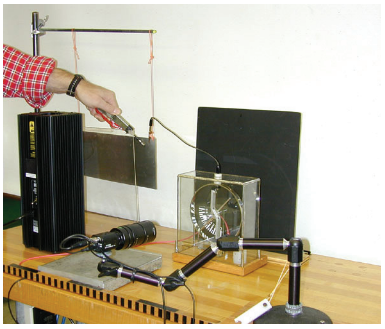
Applying this postulate to the problem of black body radiation theory, Planck obtained

$$I(\nu) = \frac{2\pi\nu^2}{c^2} \frac{h\nu}{e^{h\nu/kT} - 1} .$$

Planck's formula agrees with experimental curves. The notion that field energy is quantized rather than a continuous quantity was a new and profound addition to physics.

(*iii*) The Photoelectric Effect

A direct confirmation of the energy quantization of electromagnetic fields was provided by the phenomenon of photoelectric emission. When light of frequency ν is incident on a metal, electrons



This Demonstration shows the Production of Photo Electrons.

are emitted from the metallic surface only when $h\nu > \phi$. When $h\nu > \phi$, the electrons are emitted with a kinetic energy T, where

$$h\nu = \phi + T.$$

The energy of the photo-electron is independent of the intensity of light. The intensity of light determines merely the number (per second) of the emitted electrons. The explanation was provided by Einstein and invoked the electromagnetic field particles, photons, each carrying an energy $h\nu$. Electrons in the material are held back from the vacuum by an energy barrier of height ϕ. The impinging photon can transmit its energy $h\nu$ to an electron near the surface. If $h\nu < \phi$, this energy is insufficient to surmount the barrier and no electrons are ejected. If $h\nu > \phi$, the excess $T = h\nu - \phi$ is the kinetic energy of the electron leaving the surface.

(*iv*) The Compton Effect

The scattering of monochromatic X-rays from targets composed of light elements was studied by *Compton*. He found that the scattered radiation consists of two lines, one of the same wavelength as the incident radiation and one of slightly longer wavelength. The change in wavelength depended on the scattering angle. Compton was able to explain the measured angular dependence of wavelength shift perfectly when he assumed that individual photons of energy $E = h\nu$ and momentum $p = \frac{h\nu}{c}$ collided with individual electrons in such a way that momentum and energy were conserved. Compton's formula

Compton Effect Experiment.

$$\lambda' - \lambda = \Delta\lambda = \frac{h}{m_0 c}(1 - \cos\theta)$$

gives the wavelength of the scattered photon in terms of incident wavelength and the angle at which the scattered photon is detected. This equation was verified by Compton experimentally.

The quantity $\frac{h}{m_0 c}$ in Compton's formula has dimensions of length and is known as the Compton wavelength of the electron. Its magnitude is 2.43×10^{-12} m. This constant, which is characteristic of the X-ray scattering with wavelength shift, could not be reproduced by applying Maxwell's theory of electromagnetism to the scattering process.

(*v*) The Heat Capacity of Solids

Another experimental observation that could not be explained by ordinary physics was that of the heat capacity of solids. Classical equipartition of energy led to the prediction that the specific heat per kilo mole of solid elements should be $3R$. Dulong and Petit verified this result for many elements at high temperatures. However, the prediction failed completely in the low-termperature domain.

This disagreement between theory and experiment was resolved by Einstein and Debye who applied the Planck quantization condition

$$E = h\nu$$

to the mechanical oscillations of the lattice. The most profound effect of applying Planck's postulate to an oscillator is that its average thermal equilibrium excitation energy is no longer equal to kT (the classical value) but to

$$E = \frac{h\nu}{e^{h\nu/kT} - 1} \qquad \text{...(1)}$$

Note that $E \to kT$ for $kT >> h\nu$. The application of Eq. (1) to the heat capacity problem resulted in excellent agreement with experiment.

EXERCISE

1. What are positive rays? Describe Thomson's method for positive ray analysis. Discuss its drawbacks.

2. What is a mass spectrograph? Describe the construction of Aston's mass spectrograph with necessary theory and show how it can be used in the detection of isotopes. State the limitations of Aston's mass spectrograph.

3. Describe Bainbridge's mass spectrograph and explain how atomic masses are determined with it. Mention two uses of mass spectrograph.

4. Describe the construction, working and theory of a Dempster mass spectrograph. Can it be used for the determination of the absolute mass of an atom in practice? If not, why?

5. Give the theory of Thomson's scattering of X-rays.

6. What are the main assumptions made in the derivation of Thomson's scattering formula for X-rays? Obtain an expression for Thomson's scattering cross section. **(Delhi 1985)**

7. Explain clearly the failure of classical theory in the explanation of Compton effect.

Ans. When a beam of *X*-rays confronts matter, a certain portion of the beam is scattered in all directions (Fig. 5.13).

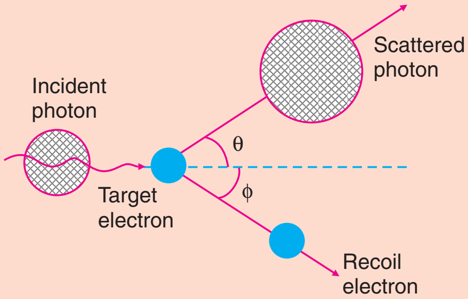


Fig. 5.13

According to the classical theory :

(*i*) the scattered *X*-ray should posses the same wavelength as the incident one

(*ii*) the scattering constant should have a constant value of 0.2 for all types of *X*-ray scattering *i.e.*, the scattering constant should be independent of the wavelength of the incident radiation

(*iii*) the distribution of intensity of *X*-rays should be symmetrical in the scattered *X*-rays.

But experiments show that:

(*i*) the scattered radiation consists of two wavelengths, one having the same wavelength as the incident one and the other a longer wavelength.

(*ii*) the scattering constant depends on the wavelength of the incident *X*-rays. It is found to decrease with decrease in the wavelength of the incident radiation.

(*iii*) the distribution of the scattered intensity is not symmetrical.

Hence, classical theory failed to explain the experimentally observed facts.

8. Show that it is not possible for a photon to transfer its entire energy to the recoil electron in a Compton scattering. **(Calcutta Hons.)**

Sol. From Art. 7.14, the expression for the maximum kinetic energy of recoil electron is

$$(E_k)_{\max} = \frac{h\nu}{1+m_0c^2/2h\nu}, \quad \text{when} \quad \theta = \angle \text{of scattering} = 180°$$

$$\angle\ h\nu, \quad \text{since } m_0c^2/2h\nu \text{ is a positive quantity.}$$

9. An *X*-ray tube is operated at 20 kV. Calculate the minimum wave length of *X*-rays emitted from it. **(Andhra University, 2010)**

Sol. Energy of electrons generating the *x*-rays = *Ve*. When the above energy is completely converted to *x*-rays, maximum frequency ν_m or minimum wavelength λ_o is obtained.

$$\therefore\ Ve = h\nu_m = \frac{hc}{\lambda_o} \qquad \therefore\ \lambda_o = \frac{hc}{Ve}$$

$$\text{Now,}\quad V = 20 \times 10^3 \text{ volt}, \quad h = 6.62 \times 10^{-34} \text{ Js}, \quad e = 1.602 \times 10^{-19} \text{ C.}$$

$$\therefore\ \lambda_o = \frac{6.62\times10^{-34}\times3\times10^8}{1.602\times10^{-19}\times20\times10^3} = 6.198 \times 10^{-11} \text{ m}$$

10. What is the mass and momentum of a photon of 1 Å wavelength?

Sol. The mass of the photon is given by

$$m = \frac{h\nu}{c^2} = \frac{h}{c\lambda} = \frac{6.6\times10^{-34}\text{ J}}{(3\times10^8\text{ m/s})\times10^{-10}\text{ m}} = \mathbf{2.2 \times 10^{-32}\ kg.}$$

The momentum of photon of mass *m* is

$$p = mc = (2.2 \times 10^{-32}) \times (3 \times 10^8) = \mathbf{6.6 \times 10^{-24}\ kg.\ m/sec.}$$

11. Calculate the average energy of an oscillator of frequency 5.6×10^{12} per second at $T = 330°$ K, treating it as (*i*) Classical oscillator, (*ii*) Planck's oscillator. **(Agra 1971)**

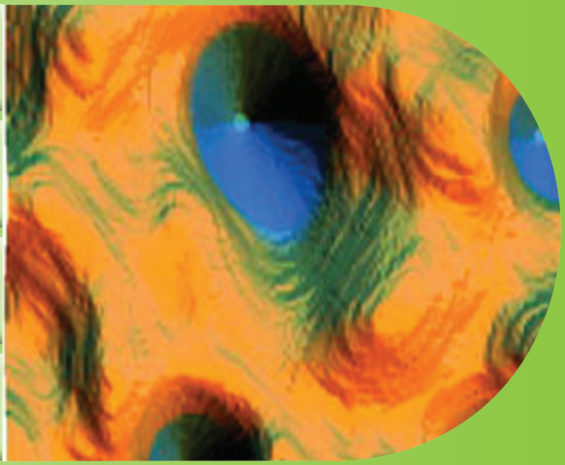
Sol. (*i*) Average energy $\bar{E}$ of a classical oscillator is

$$\bar{E} = kT = 1.380 \times 10^{-23} \times 330 = \mathbf{4.554 \times 10^{-21}\ J.}$$

(*ii*) Average energy $\bar{E}$ of a Planck's oscillator is given by

$$\bar{E} = \frac{h\nu}{\left\{\exp\left(\dfrac{h\nu}{kT}\right)-1\right\}} = \frac{6.626\times10^{-24}\times5.6\times10^{12}}{\left\{\exp\left(\dfrac{6.626\times10^{-34}\times5.6\times10^{22}}{1.380\times10^{-23}\times330}\right)-1\right\}}$$

$$= \mathbf{2.945 \times 10^{-21}\ J.}$$



CHAPTER 6

STRUCTURE OF THE ATOM

AT A GLANCE

6.1 Introduction
6.2 Rutherford's Experiments on Scattering of α-particles
6.3 Theory of α-particle Scattering
6.4 Bohr Atom Model
6.5 Effect of Nuclear Motion on Atomic Spectra
6.6 Evidences in Favour of Bohr's Theory
6.7 Correspondence Principle
6.8 Critical Potentials
6.9 Atomic Excitation
6.10 Experimental Determination of Critical Potentials
6.11 Sommerfeld's Relativistic Atom Model
6.12 The Vector Atom Model
6.13 Quantum Numbers Associated with the Vector Atom Model
6.14 Coupling Schemes
6.15 The Pauli Exclusion Principle
6.16 The Periodic Classification of Elements
6.17 Some examples of Electron Configurations with their Modern Symbolic Representations
6.18 Magnetic Dipole Moment Due to Orbital Motion of the Electron
6.19 Magnetic Dipole Moment Due to Spin
6.20 The Stern and Gerlach Experiment
6.21 Spin-Orbit Coupling
6.22 Optical Spectra
6.23 Zeeman Effect
6.24 Larmor's Theorem
6.25 Quantum Mechanical Explanation of the Normal Zeeman Effect
6.26 Anomalous Zeeman Effect
6.27 Paschen-Back Effect
6.28 Stark Effect

6.1 Introduction

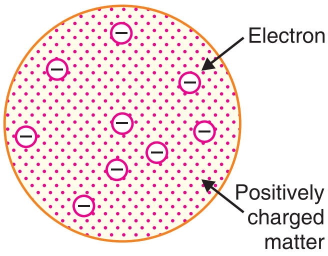


Fig. 6.1

An atom is the smallest particle of a chemical element possessing the chemical properties of the element. The first model of an atom was given by J. J. Thomson. His atom model is shown in Fig. 6.1. Thomson conceived his atomic model as a sphere of size = 10^{-10} m and of positively charged matter in which "*electrons*" were embedded. But this model could not explain all features of optical spectra of hydrogen and other elements. In 1911, Rutherford performed a number of experiments on the scattering of α-particles by a very thin gold foil. A collimated beam of alpha particles hitting a gold foil showed deflections through large-angles. Most of these particles went undeflected while a few were turned back. Thomson's model could not explain this result. Rutherford suggested that the atom consists of a central massive nucleus (dimensions of the order of 10^{-14} m) in which all the positive charge and most of the mass are concentrated. A cloud of negatively charged electrons surrounds this nucleus. Since the dimensions of the atom are of the order of 10^{-10} m, most of the space within the atom is empty. The nucleus is responsible for repelling the α-particles. The passage of undeflected α-particles was explained as due to the fact that the atom was empty all around the nucleus, the electronic mist offering no resistance to the α-particles. Thus Rutherford arrived at the concept of the nuclear atom model. Let us review Rutherford's experiment in order to make a detailed study of the atom.

6.2 Rutherford's Experiments on Scattering of α-particles

The experimental arrangement is shown in Fig. 6.2. α-particles are emitted by some radioactive material (polonium) R, kept inside a thick lead box. D is the lead screen with a small hole in it to select a very fine beam of α-particles. This well collimated beam is then allowed to fall on a thin gold foil F. While passing through the gold foil, the α-particles are scattered through different angles. A zinc sulphide screen was placed on the other side of the gold foil. This screen was movable, so as to receive the α-particles, scattered from the gold foil at angles varying from 0 to 180°. When an α-particle strikes the screen, it produces a flash of light and it is observed by the microscope M. It was found that:

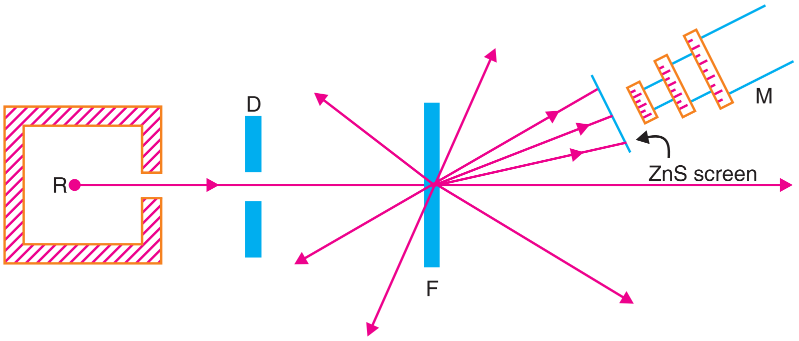


Fig. 6.2

Ernest Rutherford (1871-1937)

(1) Most of the α-particles passed through the gold foil with small deviation.

(2) There were a few particles that were scattered through large angles. Experimental results showed that one of about 8000 particles suffered angles of scattering > 90°. A few of the α-particles were even

scattered directly in the backward direction (*i.e.*, angle of scattering = 180°). Thomson's atom model failed to explain the experimental results. Therefore, Rutherford pictured an atom as follows.

Rutherford Nuclear Atom Model.

(1) The atom has a small, positively charged *nucleus.* All positive charges of an atom and most of the mass of the atom are concentrated in the nucleus. Electrons have no place inside the nucleus. The electrons revolve around this nucleus at some distance away. The dimensions of the nucleus and of the electrons are negligibly small as compared to the overall size of the atom, so that most of the volume occupied by an atom is actually an empty space. *Thus the discovery of the nucleus of the atom is due to Rutherford.*

(2) Since the electron-distribution must give stability to the atom, the electrons could not be stationary. The electrons must be revolving around the nucleus in closed orbits, so that the centrifugal force due to the rotation of the electron is balanced by the force of electrostatic attraction between the nucleus and the electrons. Thus Rutherford proposed a dynamic planetary model in which nucleus plays the role of the sun and electrons correspond to the individual planets of the solar system.

Since most of the atom is empty, most of the α-particles pass through the foil without appreciable deflection. However, if an α-particle is moving in such a direction so as to have a head-on collision with the nucleus, the α-particle will be scattered through large angles. Sometimes, when a head-on collision occurs, the α-particle is turned back. Thus the Rutherford's model of the nucleus enables us to account for very large deviations.

6.3 Theory of α-particle Scattering

We shall assume the following:

(1) The α-particle and the nucleus are point charges.

(2) The scattering is due to coulomb electrostatic repulsive forces between the α-particle and the positive charge of the nucleus (*Ze*).

(3) The nucleus is so massive compared with the α-particle that it does not move during their interaction.

(4) The α-particles do not penetrate the nuclear region and the strong interaction nuclear forces are not involved.

(5) Let *r* be the instantaneous separation between α-particle and nucleus. Owing to the variation of the electric force with $1/r^2$, the α-particle's path is a hyperbola with the nucleus at the outer focus (Fig. 6.3).

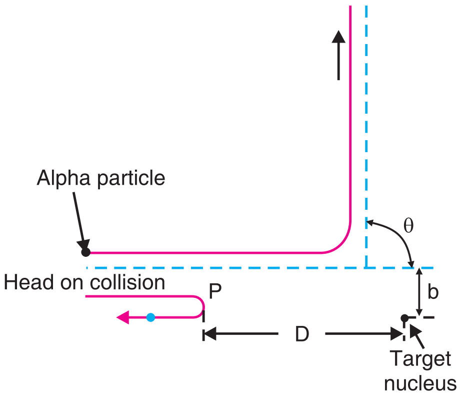


Fig. 6.3

Definitions. (1) *The impact parameter b is the minimum distance to which the α-particle would approach the nucleus if there were no forces between them.* For a head-on collision, $b = 0$. The impact parameter *b* should not be confused with the distance of closest approach *D*.

To determine *D*, consider an α-particle at a great distance from the nucleus but approaching a head-on-collision with a kinetic energy *K*. At the point *P* in Fig. 6.3, the repulsive force of the nucleus stops the approaching α-particle momentarily, and all of its K.E. is transformed into P.E. So we can write

$$K = \frac{1}{4\pi \epsilon_o} \frac{2Ze^2}{D} \quad \text{or} \quad D = \frac{1}{4\pi \epsilon_o} \frac{2Ze^2}{K}$$

(2) **The Scattering angle** θ *is the angle between the asymptotic direction of approach of the α-particle and the asymptotic direction in which it recedes.*

Relationship between b and θ. Let $\vec{p}_1$ and $\vec{p}_2$ be the linear momenta of the α-particle when it is far from the nucleus before and after the interaction respectively. Then $\frac{p_1^2}{2m} = \frac{p_2^2}{2m}$ or $p_1 = p_2 = mv$. Here v is the α-particle velocity far from the nucleus. The α-particle moves under the action of a central force $\vec{F} = \frac{1}{4\pi\epsilon_0}\left(\frac{2Ze^2}{r^2}\right)$ directed along the radius vector. From Newton's second law,

$$\Delta \vec{p} = \vec{p}_2 - \vec{p}_1 = \int_0^{\infty} \vec{F}\, dt \qquad ...(1)$$

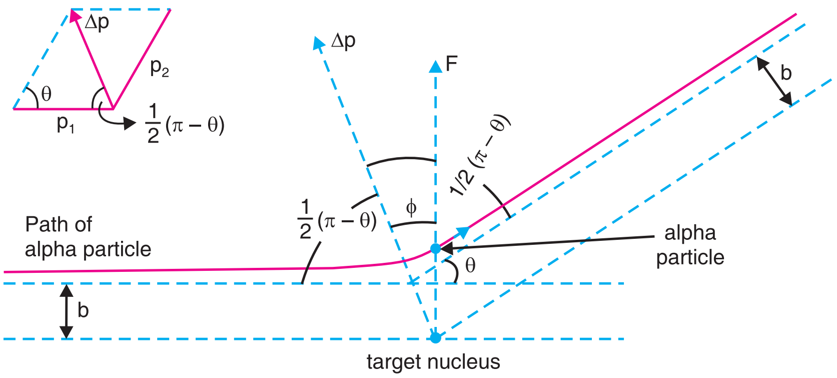


Fig. 6.4

From Fig. 6.4,

$$\frac{\Delta p}{\sin\theta} = \frac{mv}{\sin(\pi-\theta)/2}$$

or

$$\Delta p = \frac{mv\sin\theta}{\sin(\pi-\theta)/2} = \frac{mv\,2\sin(\theta/2)\cos(\theta/2)}{\cos(\theta/2)} = 2mv\sin\theta/2 \qquad ...(2)$$

Let ϕ be the instantaneous angle between $\vec{F}$ and $\Delta\vec{p}$ along the path of the α-particle. Because the impulse $\int \vec{F}\, dt$ is in the same direction as the momentum $\Delta\vec{p}$, its magnitude is $\int F\cos\phi\, dt$. Hence a scalar equation combining Equations (1) and (2) is

$$2mv\sin\frac{\theta}{2} = \int_{t=0}^{t=\infty} F\cos\phi\, dt$$

Changing variables from t to ϕ

$$2mv\sin\frac{\theta}{2} = \int_{-(\pi-\theta)/2}^{+(\pi-\theta)/2} F\cos\phi\,\frac{dt}{d\phi}\,d\phi \qquad ...(3)$$

$d\phi/dt = \omega$ is the angular velocity of the α-particle about the nucleus. The electric force exerted by the nucleus on the α-particle acts along the radius vector joining them. So there is no torque on the α-particle and its angular momentum $m\omega r^2$ is constant. Hence

$$m\omega r^2 = \text{constant} = mr^2\frac{d\phi}{dt} = mvb \quad \text{or} \quad \frac{dt}{d\phi} = \frac{r^2}{vb}$$

Substituting $F = \frac{1}{4\pi\epsilon_0}\left(\frac{2Ze^2}{r^2}\right)$ and $\frac{dt}{d\phi} = \frac{r^2}{vb}$ in Eqn. (3),

$$\frac{4\pi\epsilon_0 mv^2 b}{Ze^2}\sin\frac{\theta}{2} = \int_{-(\pi-\theta)/2}^{+(\pi-\theta)/2}\cos\phi\; d\phi = 2\cos\frac{\theta}{2}$$

or $$\cot\frac{\theta}{2} = \frac{2\pi\epsilon_0 mv^2}{Ze^2}\,b.$$

Using $K = mv^2/2$,

$$b = \frac{Ze^2}{4\pi\epsilon_0 K}\cot\frac{\theta}{2} \qquad \text{...(4)}$$

This is the relation between the impact parameter b and the scattering angle θ.

The scattering angle θ decreases with increasing impact parameter (Fig. 6.5). In terms of D,

$$b = \frac{D}{2}\cot\frac{\theta}{2} \qquad \text{...(5)}$$

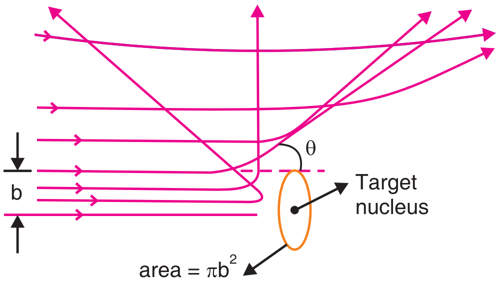


Fig. 6.5

Rutherford Scattering formula.

All α-particles that approach the nucleus with an impact parameter $\leqq b$ will be scattered by an angle $\geqq \theta$. The area around each nucleus (Fig. 6.5) with a radius equal to an impact parameter b is called the *integral cross section.* This area is

$$\sigma = \pi b^2 \qquad \text{...(6)}$$

Consider a gold foil with a surface area A and a thickness t. If n is the No. of nuclei per unit volume, the foil will contain nAt nuclei. Therefore the target area T offered by those nuclei in order to have scattering angles greater than θ and impact parameter less than b will be

$$T = nAt\sigma \qquad \text{...(7)}$$

Since A is the total target area, the fraction of α-particles experiencing scattering angles greater than θ will be

$$f = \frac{\text{target area offered by nuclei in the foil}}{\text{total target area}} = \frac{nAt\sigma}{A}$$

or $$f = nt\sigma = nt\,\pi\,b^2$$

Substituting $$b = \frac{Ze^2}{4\pi\epsilon_0 K}\cot\frac{\theta}{2},$$

$$f = \pi nt\left(\frac{Ze^2}{4\pi\epsilon_0 K}\right)^2\cot^2\frac{\theta}{2} \qquad \text{...(8)}$$

In an actual experiment, a detector measures α-particles scattered between θ and $\theta + d\theta$ (Fig. 6.6). The fraction of incident α-particles so scattered is found by differentiating Eq. (8) with respect to θ.

$$df = -\pi nt\left(\frac{Ze^2}{4\pi\epsilon_0 K}\right)^2\cot\frac{\theta}{2}\;\operatorname{cosec}^2\frac{\theta}{2}\,d\theta \qquad \text{...(9)}$$

In the experiment, a fluorescent screen was placed at a distance r from the foil. The scattered α-particles were detected by means of the scintillations they caused. Those α-particles scattered between θ and $\theta + d\theta$ reached a zone of a sphere of radius r whose width is $r\,d\theta$. The zone radius $= r\sin\theta$. So the area of the screen struck by these particles is

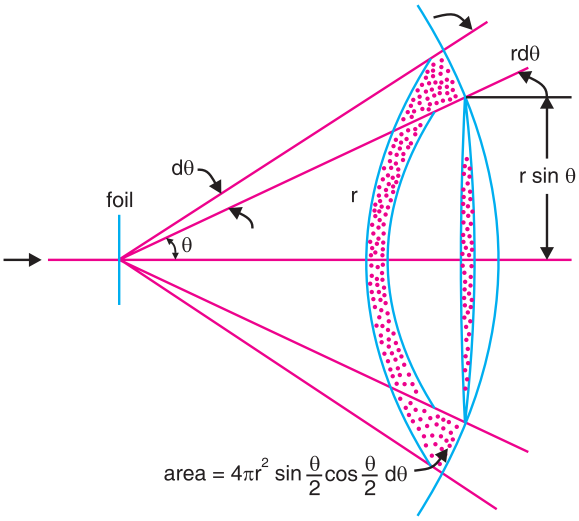

rdθ
dθ
foil
r
θ
r sin θ
area = $4\pi r^2 \sin\frac{\theta}{2}\cos\frac{\theta}{2}\, d\theta$

Fig. 6.6

$ds = (2\,\pi\, r \sin\theta)\,(r d\theta) = 2\pi r^2 \sin\theta\, d\theta = 4\pi\, r^2 \sin\theta/2 \cos\theta/2\, d\theta$.

Let N_i be the total number of α-particles directed against the target.

Then the No. scattered into $d\theta$ at $\theta = N_i\, df$

No. of particles per unit area striking the screen at θ

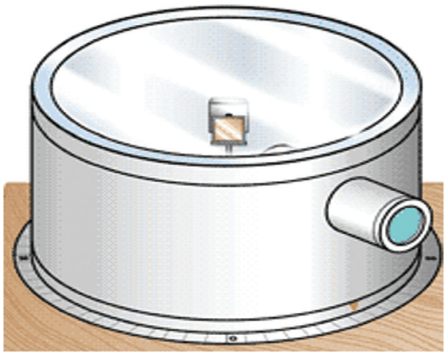
$$= N(\theta) = \frac{N_i\,|df\,|}{ds} = \frac{N_i\,\pi n t \left(\dfrac{Ze^2}{4\pi\in_0 K}\right)^2 \cot\dfrac{\theta}{2}\ \operatorname{cosec}^2\dfrac{\theta}{2}\, d\theta}{4\pi r^2 \sin\theta/2 \cos\theta/2\ d\theta}$$

$$\therefore \qquad N(\theta) = \frac{N_i\, n t\, Z^2 e^4}{(8\pi\in_0)^2\, r^2\ K^2 \sin^4(\theta/2)} \qquad \text{...(10)}$$

Eq. (10) is the Rutherford scattering formula.

Experimental Verification of the Rutherford Scattering Theory

The apparatus used by Geiger and Marsden is shown in Fig. 6.7*a*. *R* represents a radioactive substance (radon) which is the source of the α-particles. *D* is a diaphragm. *F* is a very thin foil of scattering material. *S* is a zinc sulphide screen rigidly attached to a microscope *M*. The source and foil were held fixed. The screen and microscope could be rotated. The entire apparatus was enclosed in a metal box which could be evacuated.



Geiger and Marsden Apparatus.

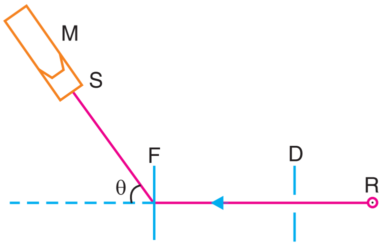


Fig. 6.7 (a)

(*i*) *Angular distribution of α-particles scattered by thin foil.* The α-particles from a radon source R passed through a diaphragm D and were scattered by the thin foil F. The α-particles that were scattered through an angle θ struck a zinc sulphide screen S. The scintillations were viewed through the microscope M. The number of particles striking the screen in a given time was measured at different angles over a range of angles from 5 to 150 degrees. Let N be the number of particles scattered per unit time through an angle θ. They found that $N \propto 1/\sin^4 (\theta/2)$.

(*ii*) *Dependence of scattering on the thickness of the foil.* The angle of scattering θ was kept constant. Foils of different thicknesses and also of different materials were used. The number of particles scattered per minute, N, is plotted as ordinates and the thickness t of the scattering foil of a given material is plotted as abscisae (Fig. 6.7*b*). Here the thickness of each foil is expressed in terms of an equivalent length of air path. The straight lines show that for any one element $N \propto t$.

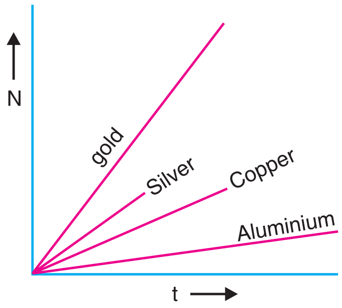


Fig. 6.7 (*b*)

(*iii*) *Dependence of scattering on the velocity of α-particles.* Velocity of the incident α-particles was changed by placing absorbing screens of mica between the source and the scattering foil. The velocities of the α-particles were determined by first finding the range R of the α-particles in air and then applying Geiger's law, $R = av^3$. Here a is a constant. It was found that Nv^4 = constant. This means $N \propto 1/v^4$ or N is inversely proportional to the square of the kinetic energy (K) of the α-particles incident on the scatterer.

(*iv*) *The dependence of the scattering on the nuclear charge.* Chadwick determined the nuclear charge from scattering experiment. The arrangement is shown in Fig. 6.7 (*c*). AA' is the scattering foil in the form of an annular ring. R is the source of α-particles. The zinc sulphide screen S was placed such that $RA = AS$. The total number of α-particles falling on the foil AA' could be determined by counting the number reaching S directly from R, since the areas of the screen S and the foil AA' were known. The α-particles that were scattered through an angle θ were counted by the scintillations they produced on the screen S. When the scattered rays were investigated, the direct rays from R were cut off completely by the thick lead plate L.

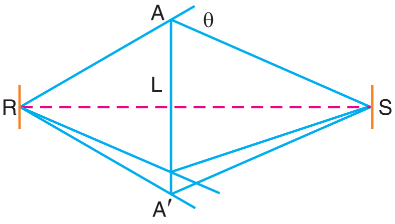


Fig. 6.7 (*c*)

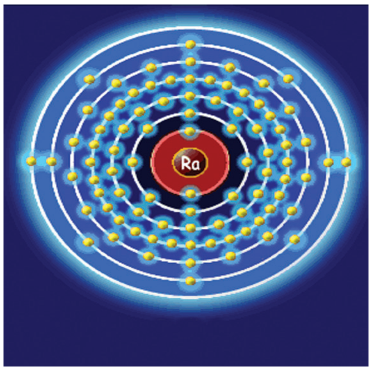
The values of Ze obtained for copper, silver and platinum foils were 29.3 e, 46.3 e and 77.4 e. These results are not precise enough to determine unique, integral values of Z. But they agree well with the values 29, 47 and 78 for the three elements, obtained by an entirely independent method. These results are in agreement with Rutherford's nuclear theory of the atom and provide the only direct measurement of the nuclear charge.

6.4 Bohr Atom Model

Bohr assumed basically Rutherford nuclear-model of the atom and tried to overcome the defects of the model. He proposed the following two postulates.

Basic Postulates. (1) *An electron cannot revolve round the nucleus in all possible orbits as suggested by the classical theory. The electron can revolve round the nucleus only in those allowed or permissible orbits for which the angular momentum of the electron is an integral multiple of $h/2\pi$.*

Here, h is the Planck's constant = 6.64×10^{-34} Js. These orbits are called stationary orbits and an electron revolving in these orbits does not radiate energy.



Bohr's Atomic Model of Radium.

Explanation. For an electron of mass m, moving with a speed v in an orbit of radius r, $L = mvr = mr^2 \omega = nh/2\pi$. n is called the *Principal quantum number*. It takes integral values, excluding zero; $n = 1, 2, 3, 4$ ....

(2) *An atom radiates energy only when an electron jumps from a stationary orbit of higher energy to one of lower energy. If the electron jumps from an initial orbit of energy E_i to a final orbit of energy E_f $(E_i > E_f)$, a photon of frequency $\nu = \frac{E_i - E_f}{h}$ is emitted.*

The Bohr formulae. Based on these postulates, Bohr derived the formulae for (*i*) the radii of the stationary orbits and (*ii*) the total energy of the electron in the orbit.

Consider an atom whose nucleus has a positive charge Ze and mass M (Fig. 6.8). For hydrogen, $Z = 1$. Let an electron of charge $(-e)$ and mass m move round the nucleus in an orbit of radius r. Since $M >> m$, the nucleus is stationary. Hence the mass of the nucleus does not come into the calculations.

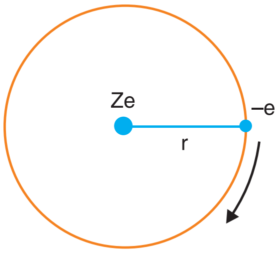


Fig. 6.8

$$\left.\begin{array}{l}\text{The electrostatic force of attraction}\\ \text{between the nucleus and the electron}\end{array}\right\} = \frac{1}{4\pi\varepsilon_0}\frac{(Ze)(-e)}{r^2} \qquad ...(1)$$

The centrifugal force on the electron $= mv^2/r$ ...(2)

$$\text{The system will be stable if } \frac{mv^2}{r} = \frac{1}{4\pi\varepsilon_0}\frac{Ze^2}{r^2} \qquad ...(3)$$

$$\text{According to Bohr's first postulate, } mvr = \frac{nh}{2\pi} \quad \text{or} \quad v = \frac{nh}{2\pi r m}$$

$$\text{or} \qquad v^2 = \frac{n^2 h^2}{4\pi^2 r^2 m^2} \qquad ...(4)$$

Substituting this value of v^2 in (3),

$$\frac{m}{r}\left(\frac{n^2h^2}{4\pi^2 r^2 m^2}\right) = \frac{1}{4\pi\varepsilon_0}\frac{Ze^2}{r^2}$$

$$\text{or} \qquad r = \frac{n^2h^2\varepsilon_0}{\pi Ze^2 m} \qquad ...(5)$$

$$\therefore \left.\begin{array}{l}\text{Radius of the } n^{th} \text{ permissible}\\ \text{orbit for hydrogen}\end{array}\right\} = r_n = \frac{n^2h^2\varepsilon_0}{\pi e^2 m} \quad (\because\ Z = 1).$$

From equation (5) we find that $r_n \propto n^2$.

The radii of the orbits are in the ratio of 1 : 4 : 9 : 16 : 25 etc.

The radius of the first orbit for hydrogen atom

$$r_1 = \frac{1^2 \times (6.625\times10^{-34})^2\ (8.854\times10^{-12})}{\pi\,(1.6\times10^{-19})^2\,(9.11\times10^{-31})}\ m = 0.053\ nm.$$

This is called the Bohr radius.

$$\therefore \qquad r_1 = 0.053\ nm\ ;\ r_2 = 2^2\ r_1 \ ....\ \text{and } r_n = n^2 r_1.$$

Calculation of total energy. The total energy of the electron in any orbit is the sum of its kinetic and potential energies. The potential energy of the electron is considered to be zero when it is at an infinite distance from the nucleus. P.E. of an electron in an orbit is given by the work done in bringing the electron from infinity to that orbit. This amount of work is obtained by integrating the electrostatic force of attraction between the nucleus and the electron from the limits ∞ to *r*.

$$\text{P.E. of the electron} = \int_{\infty}^{r} \frac{Ze^2}{4\pi\varepsilon_0 r^2}\, dr = \frac{-Ze^2}{4\pi\varepsilon_0 r} \qquad \text{...(6)}$$

$$\text{K.E. of the electron} = \frac{1}{2} mv^2 = \frac{Ze^2}{8\pi\, \varepsilon_0\, r} \qquad \text{...(7)}$$

$$\left[\because \text{ from (3)}, \quad mv^2 = \frac{Ze^2}{4\pi\, \varepsilon_0\, r}\right]$$

∴ Total energy of the electron in the n^{th} orbit = E_n = P.E. + K.E.

$$= \frac{-Ze^2}{4\pi\, \varepsilon_0\, r} + \frac{Ze^2}{8\pi\, \varepsilon_0\, r} = \frac{-Ze^2}{8\pi\, \varepsilon_0\, r}$$

Substituting the value of *r* from (5),

$$E_n = \frac{-me^4 Z^2}{8\varepsilon_0{}^2 n^2 h^2} \qquad \text{...(8)}$$

As the value of *n* increases, E_n increases. Hence, the outer orbits have greater energies than the inner orbits.

Bohr's interpretation of the Hydrogen spectrum.

If an electron jumps from an outer initial orbit n_2 of higher energy to an inner orbit n_1 of lower energy, the frequency of the radiation emitted is given by $\nu = \left(E_{n_2} - E_{n_1}\right)/h$.

$$E_{n2} = \frac{-me^4}{8\varepsilon_0{}^2 h^2} \frac{1}{n_2{}^2}$$

and

$$E_{n1} = \frac{-me^4}{8\varepsilon_0{}^2 h^2} \frac{1}{n_1{}^2} \quad [\because \ Z = 1 \text{ for hydrogen}].$$

$$\therefore \quad \nu = \frac{E_{n_2} - E_{n_1}}{h} = \frac{me^4}{8\varepsilon_0{}^2 h^3} \left\{\frac{1}{n_1{}^2} - \frac{1}{n_2{}^2}\right\} \qquad \text{...(9)}$$

The wavenumber $\bar{\nu}$ of a radiation is defined as the reciprocal of its wavelength λ in vacuum and gives the number of waves contained in unit length in vacuum. $\bar{\nu} = 1/\lambda = \nu/c$.

$$\therefore \quad \bar{\nu} = \frac{me^4}{8\varepsilon_0{}^2 ch^3} \left(\frac{1}{n_1{}^2} - \frac{1}{n_2{}^2}\right) \qquad \text{...(10)}$$

$\dfrac{me^4}{8\varepsilon_0{}^2 ch^3} = R$ and is known as *Rydberg constant.*

Substituting the values of *m*, *e*, ε_o, *c* and *h* in S.I. units,

$$R = \frac{(9.11 \times 10^{-31})\,(1.6 \times 10^{-19})^4}{8\,(8.854 \times 10^{-12})^2\,(3 \times 10^8)\,(6.625 \times 10^{-34})^3} = 1.091 \times 10^7\, m^{-1}$$

$$\therefore \quad \bar{\nu} = R\left(\frac{1}{n_1{}^2} - \frac{1}{n_2{}^2}\right) \qquad \text{...(11)}$$

Spectral series of hydrogen atom.

(1) Lyman series.

When an electron jumps from second, third, ... etc., orbits to the first orbit, we get the Lyman series which lies in the ultraviolet region. Here, $n_1 = 1$ and $n_2 = 2, 3, 4, 5$ ....

$$\therefore \qquad \bar{\nu} = R\left(\frac{1}{1^2} - \frac{1}{n^2}\right); \quad n = 2, 3, 4, 5.....$$

(2) Balmer series.

When an electron jumps from outer orbits to the second orbit, $n_1 = 2$ and $n_2 = 3, 4, 5$ ... etc.

$$\bar{\nu} = R\left(\frac{1}{2^2} - \frac{1}{n^2}\right); \quad n = 3, 4, 5.....\text{etc.}$$

This series is called Balmer series and lies in the visible region of the spectrum. The first line in the series ($n = 3$) is called the H_α line, the second ($n = 4$) the H_β line and so on.

(3) **Paschen series.** Paschen series in the infrared region are given by $n_1 = 3$ and $n_2 = 4, 5, 6$ ..... etc.

$$\therefore \qquad \bar{\nu} = R\left(\frac{1}{3^2} - \frac{1}{n^2}\right); \quad n = 4, 5, 6, 7.....\text{etc.}$$

(4) **Brackett series.** If $n_1 = 4$ and $n_2 = 5, 6, 7$ ..... etc., we get the Brackett series.

$$\therefore \qquad \bar{\nu} = R\left(\frac{1}{4^2} - \frac{1}{n^2}\right); \quad n = 5, 6, 7.....$$

(5) **Pfund series.** If $n_1 = 5$ and $n_2 = 6, 7, 8$, ... we get Pfund series.

$$\bar{\nu} = R\left(\frac{1}{5^2} - \frac{1}{n^2}\right); \quad n = 6, 7, 8.....$$

Brackett and Pfund series lie in the very far infrared region of the hydrogen spectrum. By putting $n = \infty$ in each one of the series, we get the wavenumber of the *series limit*, *i.e.*, the last line in the series. The electron jumps giving rise to the different series in hydrogen are shown in Fig. 6.9.

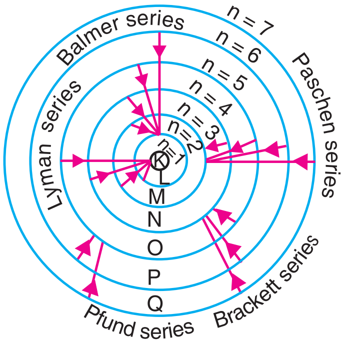


Fig. 6.9

The energy-level diagram. The equation

$$E_n = \frac{-me^4Z^2}{8\varepsilon_0^2 n^2 h^2}$$

can be diagrammatically represented. Then, it is called the energy-level diagram. Substituting the values of m, e, ε_0 and h we get for hydrogen ($Z = 1$)

$$\frac{me^4Z^2}{8\varepsilon_0^2 h^2} = \frac{(9.11 \times 10^{-31})\,(1.6 \times 10^{-19})^4\,(1^2)}{8(8.854 \times 10^{-12})^2\,(6.625 \times 10^{-34})^2} = 21.76 \times 10^{-19}\ \text{J}$$

$$= \frac{21.76 \times 10^{-19}}{1.6 \times 10^{-19}}\ \text{eV} = 13.6\ \text{eV}.$$

$$\therefore \qquad E_n = (-13.6/n^2)\ \text{eV} \quad \text{for } n = 1, 2, 3, ...$$

The lowest energy level E_1 is called the *normal* or the *ground state* of the atom and the higher energy levels E_2, E_3, E_4 ... are called the excited states. As n increases, E_n increases. As n increases, the energy levels crowd and tend to form a continuum.

In the energy-level diagram, the discrete energy states are represented by horizontal lines, and the electronic jumps between these states by vertical lines (Fig. 6.10). Fig. 6.10 shows schematically how spectral lines are related to atomic energy levels.

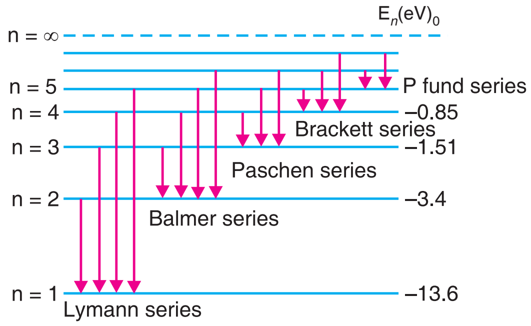


Fig. 6.10

EXAMPLE 1. *Calculate the radius and energy of the electron in the n^{th} orbit in hydrogen from the following data*:

$e = 1.6 \times 10^{-19}$ *coulomb*; $m = 9.1 \times 10^{-31}$ *kg*; $h = 6.6 \times 10^{-34}$ *joule second*; $\varepsilon_0 = 8.85 \times 10^{-12}$ *farad/metre and* $c = 3 \times 10^8$ ms^{-1}.

SOL. Radius of the n^{th} orbit for hydrogen is

$$r_n = \frac{n^2h^2\varepsilon_0}{\pi e^2 m} = \frac{n^2\,(6.6\times10^{-34})^2\,(8.85\times10^{-12})}{\pi(1.6\times10^{-19})^2\,(9.1\times10^{-31})}$$

$$\therefore \quad r_n = 0.527 \times 10^{-10}\, n^2 \text{ metre} = 0.527\, n^2 \text{ Å}.$$

$$\left.\begin{array}{l}\text{Total energy of the}\\ \text{electron in the } n^{th} \text{ orbit}\end{array}\right\} = E_n = \frac{-me^4}{8\varepsilon_0^{\,2}\, n^2h^2}$$

$$E_n = \frac{-(9.1\times10^{-31})(1.6\times10^{-19})^4}{8(8.85\times10^{-12})^2\, n^2\,(6.6\times10^{-34})^2} = \frac{-21.85\times10^{-19}}{n^2} \text{ J}$$

$$\therefore \quad E_n = \frac{-21.85\times10^{-19}}{n^2(1.6\times10^{-19})} \text{ eV} = \frac{-13.66}{n^2} \text{ eV}.$$

EXAMPLE 2. *Calculate* (*a*) *the wavenumber* (*b*) *the wavelength and* (*c*) *frequency of the* H_α *line of Hydrogen, assuming that the nucleus has infinite mass. Find the wavelength of the Balmer series limit.*

SOL. H_α line is emitted when the electron jumps from the 3rd orbit ($n_2 = 3$) to the second orbit ($n_1 = 2$).

We have $$\bar{\nu} = R\left(\frac{1}{n_1^{\,2}} - \frac{1}{n_2^{\,2}}\right)$$

Here, $$n_2 = 3,\ n_1 = 2 \text{ and } R = 1.097 \times 10^7 \text{ m}^{-1}$$

$$\therefore \quad \left.\begin{array}{l}\text{Wave number of}\\ \text{the } H_\alpha \text{ line}\end{array}\right\} = \bar{\nu} = 1.097\times10^7\left(\frac{1}{4}-\frac{1}{9}\right) = 0.1524\times10^7 \text{ m}^{-1}$$

$$\left.\begin{array}{l}\text{Wave length of}\\ \text{the } H_\alpha \text{ line}\end{array}\right\} = \lambda = \frac{1}{\bar{\nu}} = \frac{1}{0.1524\times10^7} \text{ m} = 6.561\times10^{-7} \text{ m}$$

$$\left.\begin{array}{l}\text{Frequency of}\\ \text{the } H_\alpha \text{ line}\end{array}\right\} = \nu = \frac{c}{\lambda} = \frac{3\times10^8}{6.561\times10^{-7}} = 4.561\times10^{14}\ \text{Hz}$$

By putting $n_2 = \infty$ and $n_1 = 2$, we get the wave number of the Balmer series limit.

$$\bar{\nu}_{\text{limit}} = R\left(\frac{1}{2^2}-\frac{1}{\infty^2}\right) = 1.097\times10^7\times\frac{1}{4}$$

∴ Wavelength of the Balmer series limit

$$\lambda_{\text{limit}} = \frac{1}{\bar{\nu}_{\text{limit}}} = \frac{4}{1.097\times10^7}\ \text{m} = 3646\ \text{Å}$$

EXAMPLE 3. *Show that the velocity of the electron in the first Bohr orbit is* (1/137) *c where c is the velocity of light.*

SOL. Let v be the speed of the electron (mass m, charge e) in the n^{th} orbit whose radius is r. The charge on hydrogen nucleus is $+e$. The centripetal force is provided by the electrostatic attraction.

$$\therefore \quad \frac{mv^2}{r} = \frac{1}{4\pi\varepsilon_0}\frac{e^2}{r^2} \qquad ...(i)$$

According to Bohr's first postulate, $mvr = \dfrac{nh}{2\pi}$, ...(*ii*)

Dividing Eq. (*i*) by Eq. (*ii*), $v = \dfrac{e^2}{2nh\varepsilon_0}$

The velocity in the first orbit ($n = 1$) is given by

$$v_1 = \frac{(1.6\times10^{-19})^2}{2\times1\times(6.6\times10^{-34})\,(8.854\times10^{-12})} = 2.19\times10^6\ \text{ms}^{-1}$$

$$\therefore \quad \frac{v_1}{c} = \frac{2.19\times10^6}{3\times10^8} = \frac{1}{137}.$$

EXAMPLE 4. *Using Bohr's theory of hydrogen spectrum, calculate the circumference of the first Bohr's orbit and the linear momentum of the electron moving in that orbit. Find out the de-Broglie wavelength of this electron and show how it is related to the circumference of the orbit.*

SOL. Circumference of the first Bohr orbit is

$$2\pi r = 2\pi\,(0.527\times10^{-10})\ \text{m} = 3.371\times10^{-10}\ \text{m}.$$

Linear momentum of the electron moving in the first orbit

$$mv = (9.1\times10^{-31})\,(2.19\times10^6) = 2.008\times10^{-24}\ \text{kg ms}^{-1}.$$

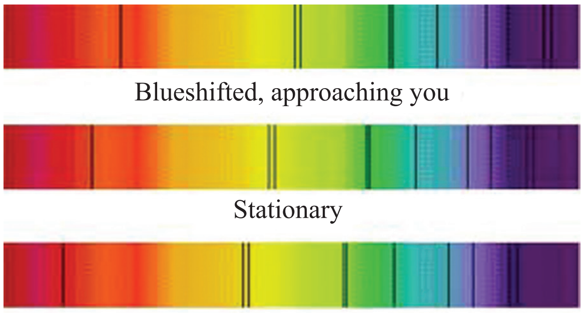
The de-Broglie wavelength of this electron is

$$\lambda = \frac{h}{mv} = \frac{6.6\times10^{-34}}{2.008\times10^{-24}} = 3.287\times10^{-10}\ \text{m}.$$

The orbit of the electron in a hydrogen atom corresponds to a complete electron de-Broglie wave joined on itself.

EXAMPLE 5. *Explain why in the absorption spectrum of hydrogen, absorption lines of only Lyman series appear.*

SOL. When a beam of white light (containing all wavelengths) is passed through hydrogen gas, and then sent into a spectrograph, a set of dark lines is obtained. In terms of quantum theory, the incident light is a beam of quanta (photons) of all sorts of energies. Now, according to Bohr theory, the hydrogen atoms absorb only those quanta whose energies correspond to transitions between its discrete energy levels. The resulting excited hydrogen atoms re-radiate the absorbed energy almost at once but these photons come off in random directions with only a few in the same direction as the original beam of white light. The dark lines in the absorption spectrum are therefore never completely black, but only appear so by contrast with the bright background. Obviously the absorption lines will have exactly the same frequencies as the emission lines.



Absorption Spectrum of Hydrogen.

Now it is found that all the emission lines of hydrogen spectrum do not appear in the absorption spectrum. The reason is that normally the atom is always in the ground state $n = 1$. Therefore, absorption transitions can only occur from $n = 1$ to $n > 1$. Hence lines of only the Lyman series can appear in absorption spectrum. Balmer lines require transitions from $n = 2$ to $n > 2$. Hence to obtain Balmer series in absorption, the atom must initially be in the state $n = 2$. Since atoms are usually in the ground state, Balmer lines are not obtained in absorption.

EXAMPLE 6. *How many revolutions does an electron in the $n = 2$ state of a hydrogen atom make before dropping to the $n = 1$ state? (The average lifetime of an excited state is about 10^{-8} s).*

SOL. Let the electron execute N revolutions before jumping from $n = 2$ to $n = 1$ orbit. Then,

$$N = \left(\frac{v_n}{2\pi r_n}\right) \times t$$

Here, v_n = velocity of electron in the nth orbit

r_n = radius to the nth Bohr orbit

t = lifetime of the electron in the excited state.

The number of revolutions of the electron in the orbit per second is

$$f = \frac{v_n}{2\pi r_n} = \frac{nh/(2\pi m r_n)}{2\pi r_n} = \frac{nh}{4\pi^2 m r_n^2} = \frac{nh}{4\pi^2 m\left(\dfrac{n^2 h^2 \varepsilon_0}{\pi m e^2}\right)^2}$$

$$f = \frac{me^4}{4\varepsilon_0^2 n^3 h^3} = \frac{(9.1\times10^{-31})\times(1.6\times10^{-19})^4}{4\times(8.854\times10^{-12})^2\times2^3\times(6.625\times10^{-34})^3} = 8.16\times10^{14}$$

$$(\because\ n = 2)$$

$$N = (8.16 \times 10^{14}) \times 10^{-8} = 8.16 \times 10^{6} \text{ revolutions.}$$

EXAMPLE 7. *Find the wavelength of the spectral line that corresponds to a transition in hydrogen from the $n = 10$ state to the ground state. In what part of the spectrum is this?*

SOL. Since the atom drops to the ground state ($n = 1$), the photon emitted belongs to the Lyman series. It lies in the ultraviolet region. The wavelengths of the spectral lines in this series are given by

$$\frac{1}{\lambda} = \bar{\nu} = R\left(\frac{1}{1^2} - \frac{1}{n^2}\right);\quad n = 2, 3, 4.....$$

$R = 1.097 \times 10^{7}$ m^{-1}. For $n = 10$, we have

$$\frac{1}{\lambda} = (1.097 \times 10^7)\left(1 - \frac{1}{100}\right)\text{m}^{-1}.$$

$$\therefore \qquad \lambda = 92.1 \times 10^{-9}\ \text{m} = 92.1\ \text{nm}.$$

6.5 Effect of Nuclear Motion on Atomic Spectra

In Bohr theory, we have assumed that the nucleus remains fixed at the centre of the circular orbit while the electron revolves round it. This is true if the nucleus has infinite mass. If the nucles has a mass M that is not infinite, both the nucleus and its orbital electron revolve around a common centre of mass with same angular velocity ω. Let N and e represent (Fig. 6.11) the nucleus and the electron of mass M and m respectively. Now both of them are rotating about their common centre of mass C, the nucleus moving in a circle of radius r_1 and the electron in a circle of radius r_2.

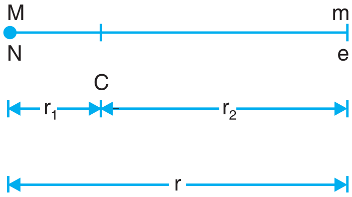


Fig. 6.11

Now according to centre of mass theory,

$$Mr_1 = mr_2 \qquad \text{...(1)}$$

Let r represent the distance between the nucleus and electron.

Then, $$r = r_1 + r_2 \qquad \text{...(2)}$$

$$\therefore \qquad r_1 = r - r_2 = r - \frac{Mr_1}{m}$$

or $$r_1\left(1 + \frac{M}{m}\right) = r$$

or $$r_1 = \left(\frac{m}{M+m}\right)r \qquad \text{...(3)}$$

Similarly $$r_2 = \left(\frac{M}{M+m}\right)r \qquad \text{...(4)}$$

The total angular momentum of the atom about the centre of mass is

$$\begin{aligned} L &= Mr_1^{\,2}\,\omega + mr_2^{\,2}\,\omega \\ &= M\left(\frac{mr}{M+m}\right)^2\omega + m\left(\frac{Mr}{M+m}\right)^2\omega \\ &= \left(\frac{Mm}{M+m}\right)r^2\,\omega \\ &= \mu\, r^2\,\omega \qquad \text{...(5)} \end{aligned}$$

where $\mu = \dfrac{Mm}{M+m} = \dfrac{m}{1+(m/M)}$ is called the reduced mass of the electron.

According to Bohr's first postulate,

$$\mu\, r^2\,\omega = n\,(h/2\pi) \qquad \text{...(6)}$$

In the absence of nuclear motion, the corresponding equation was

$$mr^2\,\omega = n(h/2\pi) \qquad \text{...(7)}$$

Eq. (6) is identical to Eq. (7), except that m has been replaced by μ. Therefore to take into account the finite mass of the nucleus we must replace the mass of the electron (m) in all the Bohr formulae by the reduced mass (μ).

Equation for the wavenumbers of spectral lines of the atom.

The energy-levels of the hydrogen atom become

$$E_n = -\frac{\mu e^4 Z^2}{8\varepsilon_0{}^2 h^2}\left(\frac{1}{n^2}\right) \qquad ...(8)$$

When an electron jumps from the Bohr orbit n_2 to the Bohr orbit n_1, the frequency of the radiation emitted is

$$\nu = \frac{\mu e^4 Z^2}{8\varepsilon_0{}^2 h^3}\left(\frac{1}{n_1{}^2}-\frac{1}{n_2{}^2}\right)$$

or

$$\bar{\nu} = \frac{\mu e^4 Z^2}{8\varepsilon_0^2 ch^3}\left[\frac{1}{n_1^2}-\frac{1}{n_2^2}\right] \qquad ...(9)$$

The Rydberg constant for any element is given by,

$$R_Z = \frac{\mu e^4}{8\varepsilon_0^2 ch^3} = \frac{me^4}{8\varepsilon_0^2\, ch^3}\times\frac{M_z}{(M_z+m)}$$

$$= R_\infty \times \frac{1}{1+m/M_z} \qquad ...(10)$$

where $R_\infty = \frac{me^4}{8\varepsilon_0^2\, ch^3}$ = The value of the Rydberg constant when $M_Z = \infty$ *i.e.*, when the nucleus is at rest and M_Z = Mass of the nucleus of the element of atomic number Z. The constant R_∞ can be estimated from the spectroscopically observed value of the Rydberg constant for hydrogen R_H = 10067770 m^{-1} and the known value of $m/M_H = 1/1840$ using the relation $R_H = R_\infty\,\{1/(1 + m/M_H)\}$.

The value of R_∞ thus found is 10973740 m^{-1}.

Equation (9) can be written as

$$\bar{\nu} = Z^2 R_Z\left(\frac{1}{n_1{}^2}-\frac{1}{n_2{}^2}\right) \qquad ...(11)$$

6.6 Evidences in Favour of Bohr's Theory

(1) **The ratio of mass of electron to the mass of proton.**

The Rydberg constants for hydrogen and helium are

$$R_H = \frac{R_\infty}{1+m/M_H} \text{ and } R_{He} = \frac{R_\infty}{1+m/M_{He}}$$

$$\therefore \quad \frac{R_{He}}{R_H} = \frac{1+m/M_H}{1+m/M_{He}}$$

But $M_{He} \approx 4M_H$.

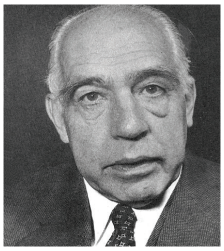
$$\therefore \quad \frac{R_{He}}{R_H} = \frac{1+m/M_H}{1+m/4M_H}$$

or

$$R_{He} + \frac{R_{He}}{4}\frac{m}{M_H} = R_H + R_H\frac{m}{M_H}$$

or

$$R_{He} - R_H = \frac{m}{M_H} = \left(R_H - \frac{R_{He}}{4}\right)$$



Niels Bohr
(1885 – 1962)

$$\therefore \qquad \frac{m}{M_H} = \frac{R_{He} - R_H}{R_H - R_{He}/4}$$

But from spectroscopic data, $R_{He} = 10972240\ m^{-1}$ and $R_H = 10967770\ m^{-1}$

$$\therefore \qquad \frac{m}{M_H} \approx \frac{1}{1837}.$$

This value is in excellent agreement with the value obtained by other methods.

(2) **Spectrum of singly ionised helium.**

Singly ionised helium He^+ (a helium atom which has lost a single electron) resembles a hydrogen atom, except that $Z = 2$ and the nucleus is nearly four times as heavy. Putting $Z = 2$ in Eqn. (11),

$$\bar{\nu} = 4R_H\left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right).$$

Thus we see that singly-ionised helium should emit the same type of atomic spectrum as hydrogen, except that all wavenumbers are approximately four times as large. This conclusion from the theory agrees with observation except for a slight numerical discrepancy due to the fact that the reduced mass μ is slightly greater for *He* than for *H*.

(3) **The discovery of deuterium.** Deuterium is an isotope of hydrogen whose atomic mass is double that of ordinary hydrogen owing to the presence of a neutron as well as a proton in the nucleus. According to theory, for atoms of the same value of *Z*, there should be lines of slightly different wavenumbers if their nuclei have different masses. Because of the greater nuclear mass, the spectral lines of deuterium are all shifted slightly to wavelengths shorter than those of ordinary hydrogen. For example, H_α line of deuterium has a wavelength of 6561 Å, while that of hydrogen is 6563 Å. It was found by Urey and his co-workers that in the atomic spectrum of natural hydrogen, lines due to deuterium were slightly displaced from the corresponding lines due to hydrogen proper. It was also observed by them that the intensity of lines which are slightly shifted towards the short wavelength side is extremely less than the corresponding hydrogen lines. This is because the concentration of deuterium in natural hydrogen is only one atom in 5000.

EXAMPLE. *Determine the separation of the first line of the Balmer series in a spectrum of ordinary hydrogen and tritium (mass number 3). Given that* $R_\infty = 10973740\ m^{-1}$.

SOL. Rydberg constant for ordinary hydrogen atom is

$$R_H = \frac{R_\infty}{(1 + m/M)} \qquad \text{...(1)}$$

Rydberg constant for tritium is

$$R_T = \frac{R_\infty}{(1 + m/3M)} \qquad \text{...(2)}$$

Let λ_1 and λ_2 be the wavelengths of the first line of Balmer series due to the two isotopes. Then,

$$\frac{1}{\lambda_1} = R_H\left(\frac{1}{2^2} - \frac{1}{3^2}\right) = \frac{5}{36}R_H \quad \text{or} \quad \lambda_1 = \frac{36}{5R_H} \qquad \text{... (3)}$$

$$\frac{1}{\lambda_2} = R_T\left(\frac{1}{2^2} - \frac{1}{3^2}\right) = \frac{5}{36}R_T \quad \text{or} \quad \lambda_2 = \frac{36}{5R_T} \qquad \text{... (4)}$$

$$\therefore \qquad \lambda_1 - \lambda_2 = \frac{36}{5}\left(\frac{1}{R_H} - \frac{1}{R_T}\right) = \frac{36}{5}\left[\frac{(1 + m/M)}{R_\infty} - \frac{(1 + m/3M)}{R_\infty}\right]$$

$$= \frac{36}{5R_\infty}\left[\frac{2m}{3M}\right] = \frac{36}{5\times 10973740}\times\frac{2}{3}\times\frac{1}{1836}$$

$$= 2.38 \times 10^{-10}\ m = 2.38\ \text{Å}.$$

Ritz combination principle.

Statement. *By a combination of the terms that occur in the Rydberg or Balmer formula, other relations can be obtained holding good for new lines and new series.*

By this principle, Ritz predicted new series of lines in the hydrogen spectrum before they were actually discovered.

Example. *Consider the first two lines, H_α and H_β of the Balmer series. We may represent them by*

$$\bar{\nu}_\alpha = R\left(\frac{1}{2^2}-\frac{1}{3^2}\right) \quad \text{and} \quad \bar{\nu}_\beta = R\left(\frac{1}{2^2}-\frac{1}{4^2}\right)$$

Combining these two formulae,

$$\bar{\nu}_\beta - \bar{\nu}_\alpha = R\left(\frac{1}{3^2}-\frac{1}{4^2}\right)$$

This represents a new line. It is the first line of a new series in the infrared, discovered by Paschen. Similarly, the second line of the Paschen series can be obtained by forming the difference of H_γ and H_α and so on.

Ritz combination principle may also be stated as follows : *If lines at frequencies ν_{ij} and ν_{jk} exist in a spectrum with $j > i$ and $k > j$, then there will usually be a line at ν_{ik} where $\nu_{ik} = \nu_{ij} + \nu_{jk}$.* However, not all combinations of frequencies are observed because certain selection rules operate which will be discussed later.

Example. *If lines of frequencies ν_{12} and ν_{23} can be represented as*

$$\nu_{12} = T_1 - T_2;\ \ \nu_{23} = T_2 - T_3$$

then a line of frequency ν_{13} will exist, where

$$\nu_{13} = (T_1 - T_2) + (T_2 - T_3) = T_1 - T_3$$

Alkali atomic spectra

Hydrogen and the alkali metals Li, Na, K, Rb, and Cs form group 1 of the periodic table.

All these atoms have a single *s*-electron outside a closed shell. This *s*-electron is called the *valence electron* since it determines the chemical characteristics of the atom.

In all these atoms except the hydrogen, the valence electron moves in a net field of the nucleus of positive charge $+Ze$ and the core of electrons with negative charge $-(Z-1)e$ surrounding the nucleus. These electrons act as a shield between the valence electron and the nucleus. Owing to this shielding of the attraction of the valence electron by the other electrons, the effective nuclear charge is not Ze but a lesser value Z_{eff}. The energy is then given by

$$E_n = -\frac{RhcZ_{\text{eff}}^2}{n^2} = -\frac{Rhc}{n^2}(Z-\sigma)^2$$

Here, σ is a *screening constant* which is different for the different *l*-states of a given value of *n*. The value of Z_{eff} is largest for 3*s*. Hence it is lowered much more than the 3*p* state. For a given *n*, the *S* state has less energy than the *P* state, the *P* state less energy than the *D* state, the *D* state less than the *F* state, and so on. With increasing *n,* the energy difference between states becomes less and less. For sodium the 5*d* and 5*f* levels almost coincide (See Chapter 25, Fig. 25.3).

The spin angular momentum of the valence electron combines with its orbital angular momentum and gives the total angular momentum $\sqrt{j(j+1)}\,\hbar$. If $l \neq 0$, the total angular momentum quantum number j can have the values $j = l + (1/2)$ or $j = l - (1/2)$. That is, each of the l levels ($l \neq 0$) in Fig. 25.3 splits into a doublet. If $l = 0$, j takes the only value 1/2. Therefore, the S states remain as singlet. The selection rules for the transitions are

Δn = any value, $\Delta l = \pm 1$, $\Delta j = 0, \pm 1$.

6.7 Correspondence Principle

In 1932 Niels Bohr proposed a correspondence principle.

Statement. *Any new theory in Physics must reduce to well-established corresponding classical theory when the new theory is applied to the special situation in which the less general theory is known to be valid.*

Explanation. Bohr's theory gives only the frequencies of the spectral lines and says nothing about the nature (whether polarised or not) and intensity of lines, whereas classical theory is very successful in this respect. Also, according to classical theory, the frequency of the spectral line is the same as the orbital frequency of the electron ($\nu = \omega/2\pi$). But in Bohr's theory, the frequency of the spectral line is determined by the difference in energy between two orbital states: $\nu = (E_i - E_f)/h$. But it can be shown that, for transitions between states whose quantum numbers are relatively high, the frequency of the spectral line coincides very nearly with the orbital frequency.

Let us consider an atom of effectively infinite mass. Then

$$E_n = \frac{-me^4}{8\varepsilon_0^2 n^2 h^2} \qquad \text{...(1)}$$

If n is sufficiently great, the energy change ΔE corresponding to a change of n by Δn is obtained by differentiating equation (1).

$$\Delta E_n = \frac{me^4}{4\varepsilon_0^2 h^2 n^3}\,\Delta n \qquad \text{...(2)}$$

$$\therefore \qquad \nu = \frac{\Delta E_n}{h} = \frac{me^4}{4\varepsilon_0^2 h^3 n^3}\,\Delta n \qquad \text{...(3)}$$

According to Bohr's first postulate,

$$nh/2\pi = mr^2\omega$$

$$\text{or} \qquad nh = 2\pi mr^2\omega.$$

$$\therefore \qquad \nu = \frac{me^4}{4\varepsilon_0^2 (2\pi mr^2\omega)^3}\,\Delta n$$

$$= \frac{me^4}{32\pi^3 \varepsilon_0^2 m^3 r^6 \omega^3}\,\Delta n \qquad \text{...(4)}$$

Now for the equilibrium in the orbit, we have,

$$mr\omega^2 = \frac{1}{4\pi\varepsilon_0}\frac{e^2}{r^2} \quad \text{or} \quad \frac{1}{r^3} = \frac{4\pi\varepsilon_0 m\omega^2}{e^2}$$

$$\therefore \qquad \nu = \frac{me^4}{32\pi^3 \varepsilon_0^2 m^3 \omega^3}\left(\frac{4\pi\varepsilon_0 m\omega^2}{e^2}\right)^2 \Delta n$$

or $$\nu = \frac{\omega}{2\pi}\Delta n \qquad \text{...(5)}$$

Now, if $$\Delta n = 1, \nu = \frac{\omega}{2\pi}.$$

Thus the frequency given by the quantum theory for two very large quantum numbers and separated by unity becomes identical with the orbital frequency and hence with the classical frequency. Therefore, we may conclude that the behaviour of the atom tends asymptotically to that expected from the classical theory in the region of large quantum numbers.

This correspondence principle has proved to be of great value in the computation of the intensity, polarisation and coherence of spectral radiation, as well as in the formulation of the so-called selection rules.

Special Hydrogenic Systems: Positronium, muonic atom

Example 1. *A positronium "atom" is a system that consists of a positron and an electron that orbit each other. This combination resembles the hydrogen atom. Compare the wavelengths of the spectral lines of positronium with those of ordinary hydrogen.*

Sol. The positron has the same mass m as the electron and has equal but positive charge. The reduced mass of the electron-positron atom is

$$m' = \frac{(m)(m)}{m+m} = \frac{1}{2}m$$

The energy of the electron in nth orbit of hydrogen ($Z = 1$) is given by

$$E_n = -\frac{me^4}{8\varepsilon_0^2 h^2}\left(\frac{1}{n^2}\right).$$

For the ground orbit, $n = 1$.

$$E_1 = -\frac{me^4}{8\varepsilon_0^2 h^2}$$

$$\therefore \qquad E_n = \frac{E_1}{n^2}.$$

The reduced mass of electron in hydrogen is very nearly m.

Taking nuclear motion into account, the energy of the electron in the nth orbit of the hydrogen atom is

$$E_n = -\frac{me^4}{8\varepsilon_0^2 h^2}\left(\frac{1}{n^2}\right) = \frac{E_1}{n^2}$$

The energy levels of a positronium "atom" are

$$E'_n = \left(\frac{m'}{m}\right)\frac{E_1}{n^2} = \frac{E_1}{2n^2}$$

This means that the Rydberg constant for positronium is half as large as it is for ordinary hydrogen. As a result, the wavelengths in the positronium spectral lines are all twice those of the corresponding lines in the hydrogen spectrum.

Example 2. *A proton captures a μ^- meson (charge –e, mass = 207 m, where m is mass of electron) to form a muonic atom. (a) Find the radius of the first Bohr orbit of this atom. (b) Find the ionization energy of the atom.*

Sol. The reduced mass of the system is

$$m' = \frac{(207\,m)(1836\,m)}{207\,m + 1836\,m} = 186\,m\,\cdot$$

From Bohr theory, the radius of the first orbit ($n = 1$) of a hydrogen atom for $Z = 1$, is given by

$$r_1 = \frac{h^2\varepsilon_0}{\pi m e^2}$$

Here, $r_1 = a_0 = 5.29 \times 10^{-11}$m.

Hence the radius r' that corresponds to the reduced mass m' is

$$r'_1 = \left(\frac{m}{m'}\right) r_1 = \left(\frac{m}{186\,m}\right) a_0 = 2.85 \times 10^{-13}\,m.$$

The muon is 186 times closer to the proton than an electron would be.

(*b*) The energy levels of the atom corrected for nuclear motion are

$$E'_n = -\frac{m'e^4}{8\varepsilon_0^2 h^2}\left(\frac{1}{n^2}\right) = \left(\frac{m'}{m}\right)\left(\frac{E_1}{n^2}\right)$$

With $n = 1$ and $E_1 = -13.6$ eV, we have

$$E'_1 = \left(\frac{m'}{m}\right) E_1 = 186\ E_1 = -2.53 \times 10^3\ \text{eV} = -2.53\,\text{keV}$$

The ionization energy is therefore 2.53 keV, 186 times that for an ordinary hydrogen atom.

6.8 Critical Potentials

Definitions. (1) **Critical potential.** *The least energy, expressed in electron volts, required to excite a free neutral atom from its ground state to a higher state is called a critical potential of the atom.*

It is usual to distinguish two kinds of critical potentials, *viz.*, excitation potential (radiation potential or resonance potential) and ionisation potential.

(2) **Excitation potential.** *The energy in electron volts required to raise an atom from its normal state into an excited state is called excitation potential of the state.*

(3) **Ionisation potential** *of the atom may be defined as the energy required to remove an electron from a given orbit to an infinite distance from the nuclus.*

Example. *We have seen that the energy associated with an electron in the nth orbit of the hydrogen atom is* $E_n = -13.6/n^2$ *volts. Thus the energies of the first, second, third ... orbits are respectively – 13.6, – 3.4, – 1.51, ... 0 electron volts. The energy required to raise the atom from the ground state (n = 1) to the first excited state is 13.6 – 3.4 = 10.2 eV. The energy required to raise it to the second excited state is (13.6 – 1.51) = 12.09 eV and so on. The ionisation potential = 13.6 eV. It is clear that 10.2, 12.09 are excitation potentials, while 13.6 volts is the ionisation potential of the hydrogen atom.*

6.9 Atomic Excitation

There are two main ways of exciting an atom.

(*i*) Excitation by Collision. One method of exciting the atoms is to bombard them with electrons having enough kinetic energy. To produce a luminous discharge in a rarefied gas, an electric field is established that accelerates electrons until their kinetic energies are sufficient to excite atoms they collide with. The electron gives up all or part of its energy in exciting the atom. The atom then emits a photon in returning to its ground state. Neon signs and mercury vapour lamps are familiar examples of how a strong electric field applied between electrodes in a gas-filled tube leads to the emission of the characteristic spectral radiation of that gas. We get reddish light in the case of neon and bluish light in the case of mercury vapour.

(*ii*) Irradiation of atoms with light. Atoms can be excited by energy supplied in the form of light. An atom absorbs a photon of light whose energy is just the right amount to raise the atom to a higher energy level. For example, a photon of wavelength 121.7 nm is emitted when a hydrogen atom in the $n = 2$ state drops to the $n = 1$ state. Absorbing a photon of wavelength 121.7 nm by a hydrogen atom initially in the $n = 1$ state will therefore bring it up to the $n = 2$ state. This process explains the origin of absorption spectra.

When white light, which contains all wavelengths, is passed through hydrogen gas, photons of those wavelengths that correspond to transitions between energy levels are absorbed. The resulting excited hydrogen atoms radiate their excitation energy almost at once. But these photons come off in random directions with only a few in the same direction as the original beam of white light. The dark lines in an absorption spectrum are therefore never completely black, but only appear so by contrast with the bright background. We expect the absorption spectrum of any element to be identical with its emission spectrum and this agrees with observations. For example, in the case of sodium, the two black absorption lines occupy exactly the same position in the spectrum as the two yellow D_1 and D_2 emission lines.

6.10 Experimental Determination of Critical Potentials

(1) Franck and Hertz's method

Experimental arrangement. The gas of the element under study (Franck and Hertz had studied mercury vapour) is filled in a glass tube (T) at a pressure of about 1 mm of mercury (Fig. 6.12). Electrons are produced by heating the filament (F) by a low tension battery (B). These electrons are accelerated towards a grid G by the potential difference V between F and G. V can be varied between 0 and 60V by a potentiometer arrangement. P is the collector plate which is kept at a slightly negative potential (V'), about 0.5 V, with respect to G. Thus, only those electrons from G can go to P which have kinetic energies greater than this potential difference. The milliammeter A measures the plate current.

Franck and Hertz's Experimental Arrangement.

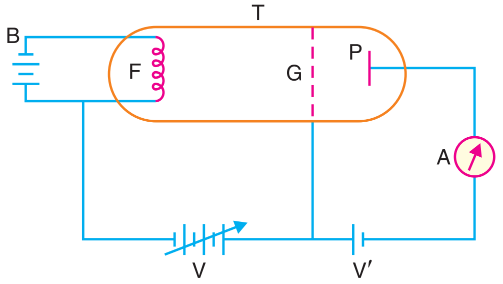


Fig. 6.12

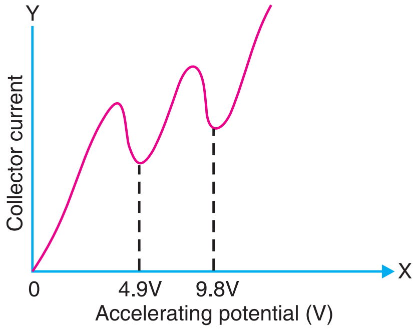


Fig. 6.13

Keeping V' constant, V is gradually increased in small steps from zero upwards. A plot of the collector current against the accelerating potential V is shown in Fig. 6.13. From this graph, we see that there is no collector current for $V < 0.5\ V$. Above this, the collector current increases continuously. When the accelerating P.D. reaches a value 4.9 volt, the current suddenly dips to a minimum. Again when the P.D. is gradually increased above 4.9 volt, the current gradually increases, till another maximum is reached when the P.D. is just below 9.8 volts. At 9.8 volts, the current again dips steeply to a minimum. Similarly, a significant decrease in the collector current occurs each time the accelerating potential is increased by approximately $5V$.

Explanation of the graph. Consider the case of a beam of slow electrons travelling through mercury vapour at low pressure. If the electrons suffer no energy loss due to collision with the atoms of the gas, they will reach the collector plate. On increasing the accelerating potential, collector current increases. As V reaches the value of a critical potential (4.9 volt), an electron acquires 4.9 eV of energy on reaching G. The electron loses all its energy in an inelastic collision with a mercury atom. Thus the electron is left with no energy to reach P. Consequently the collector current drops abruptly. This suggests that the atom has absorbed this 4.9 eV energy to raise it from the ground state to a state of higher energy.

This current dipping to a minimum at 4.9 volts does not reach zero because statistically some electrons may succeed in reaching the collector plate, avoiding an inelastic collision with a mercury atom. Such electrons record minimum current on the plate. Then as V is further increased, the electrons start reaching P after the inelastic collisions, if they are left with energy to overcome V'. Hence, the plate current again increases. At $V = 9.8$ volts, the current again dips to a second minimum. This can be explained if an electron of 9.8 eV energy suffers two consecutive inelastic collisions with different mercury atoms before it reaches the plate. Such an electron excites both the mercury atoms to their first excited states, losing 4.9 eV energy in each collision. This explains the second minimum.

Each time there is an inelastic collision, the mercury atoms will be excited and return to the ground state by the emission of photons. By using spectroscopic techniques, the wavelength of the radiation coming from the tube was found to be 2536 Å corresponding to transitions from the first excited state to the ground state. Let us see what is the energy of a photon of this wavelength.

$$E = h\nu = \frac{hc}{\lambda} = \frac{(6.63 \times 10^{-34})\,(3 \times 10^{8})}{2536 \times 10^{-10}}\ \text{J}$$

$$= \frac{(6.63 \times 10^{-34})\,(3 \times 10^{8})}{(2536 \times 10^{-10})\,(1.6 \times 10^{-19})}\ eV = 4.9\ eV \cdot$$

Thus, this experiment shows that the energy lost by the electron in its inelastic collision with the mercury atom reappears as a quantum of energy of wavelength hc/E. This experiment shows in a very convincing way the existence of discrete energy levels in the mercury atom. The main drawback of this experiment is that one cannot distinguish between the excitation and ionisation potentials.

EXAMPLE 1. *Sodium has first two excited states at* 2.1 *and* 3.7 *eV. In a Franck-Hertz experiment, electrons of energy* 4.7 *eV are fired in sodium gas.* (*i*) *Deduce the possible energy values of the electrons received at the collector.* (*ii*) *Deduce the wavelength of spectral lines expected.*

SOL. (*i*) The electron loses energy in its inelastic collision with the sodium atom. In collisions, the electrons may lose energy 0, or 2.1 eV, or 3.7 eV, or (2 × 2.1) eV, or (2 × 3.7) eV or (2.1 + 3.7) eV. Out of these, the values below 4.7 eV are 0, 2.1, 3.7 and 4.2 eV. Hence the possible energies of the received electrons are: (4.7 – 0), (4.7 – 2.1), (4.7 – 3.7) and (4.7 – 4.2) eV. Thus $E = 4.7, 2.6, 1.0$ and 0.5 eV.

(*ii*) The excited states are 2.1 eV and 3.7 eV. The spectral lines will therefore correspond to 3.7 eV, 2.1 eV and (3.7 – 2.1) = 1.6 eV.

Now $$E = \frac{hc}{\lambda} \quad \text{or} \quad \lambda = \frac{hc}{E}$$

$$\therefore \qquad \lambda_1 = \frac{(6.6\times10^{-34})\,(3\times10^{8})}{3.7\times(1.6\times10^{-19})} = 3.3\times10^{-7}\,m$$

Similarly, $$\lambda_2 = 5.8\times10^{-7}m \text{ and } \lambda_3 = 7.7\times10^{-7}m.$$

EXAMPLE 2. *Calculate (i) ionisation potential (ii) first excitation potential of the hydrogen atom taking* $h = 6.62\times10^{-34}$ *Js,* $e = 1.6\times10^{-19}$ *C and* $m = 9.1\times10^{-31}$ *kg.*

SOL. The ionisation potential is numerically equal to the energy of the electron in the first orbit.

$$E_1 = \frac{-me^4}{8\varepsilon_0^2 h^2} = \frac{-(9.1\times10^{-31})\,(1.6\times10^{-19})^4}{8\,(8.854\times10^{-12})^2\,(6.62\times10^{-34})^2}$$

$$= -21.7\times10^{-19}\,J.$$

$$= \frac{-21.7\times10^{-19}}{1.6\times10^{-19}}\,eV = -13.6\,eV.$$

Hence the ionisation potential of the hydrogen atom is 13.6 *V*.

(*ii*) The first excitation potential of the atom is the energy required to shift the electron from $n = 1$ to $n = 2$ orbit, *i.e.*, $E_1 \sim E_2$.

$$E_1 = -13.6\,eV.$$

Since $$E_n \propto \frac{1}{n^2}, \quad E_2 = \frac{E_1}{4} = -3.4\,eV.$$

$$\therefore \qquad E_1 \sim E_2 = 13.6 - 3.4 = 10.2\,eV.$$

Hence the first excitation potential of the hydrogen atom is 10.2 *V*.

(2) Davis and Goucher's method

A glass bulb, containing the gas under study (mercury), has a filament *F*, cathode *C*, grids G_1 and G_2 and a plate *P* as shown in Fig. 6.14, very similar to a tetrode. An accelerating potential V_a is applied between *C* and G_1. The electrons from the filament *F* are accelerated by the P.D. V_a and enter the collision region between the two grids G_1 and G_2. In the collision region G_1 G_2, the electrons experience a retarding P.D. V_r just greater than V_a. A small P.D. *V* is maintained between G_2 and *P*.

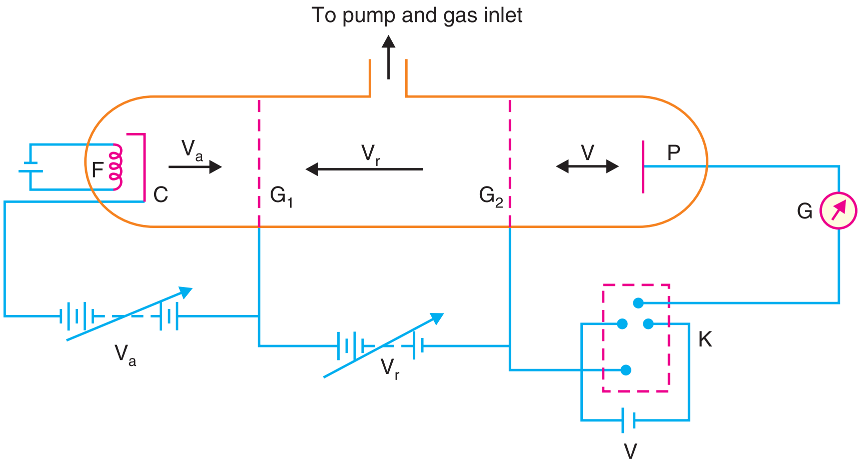


Fig. 6.14

This potential can be reversed with the help of reversing key K. The grid G_2 and the metal plate P are made of photosensitive material. The current between G_2 and P is indicated by a galvanometer G. Davis and Goucher, obtained for mercury vapour, curves of the kind shown in Fig. 6.15 by plotting the galvanometer current against the accelerating potentials V_a.

The accelerating potential V_a is gradually increased from zero. As V_r is always maintained just greater than V_a, no electrons are able to reach the plate whatever be the direction of V. The galvanometer current is, therefore, zero. When V_a becomes equal to an excitation potential, the electrons excite gas atoms in the region G_1 G_2. In the case of mercury, for example, this occurs when V_a approaches $4.7V$. The mercury atoms excited by inelastic collisions with the electrons of energy 4.7 eV, return to the ground state with the emission of spectral radiation of wavelength 2536 Å. This radiation falling on G_2 and P releases photoelectrons. Let the reversing key be such that G_2 is negative with respect to P. The photoelectrons released from P cannot move towards G_2, while the photoelectrons released from G_2 move towards P. Therefore, the galvanometer indicates a current. This is the state corresponding to the point *s* on the curve.

When V_a is increased beyond $4.7V$, due to deeper penetration of the accelerated electrons into the collision region, more and more atoms in this region get excited. Consequently, the photoelectric current increases. If the potential between G_2 and P is reversed and the experiment repeated, photoelectrons can move from P to G_2 and not from G_2 to P. This results in a current in the galvanometer in a direction opposite to that previously obtained. When V_a reaches 6.7 volts, there is a rapid increase in the photoelectric current. This is due to the emission of the spectral radiation of wavelength 1849 Å by the mercury atoms excited by inelastic collisions with electrons. As before, this photoelectric current is reversed on reversing V.

In Fig. 6.15 the curve *sqr* above the X-axis is obtained when G_2 is positive, while the curve *sq'r'* below the X-axis when G_2 is negative with respect to P. The next abrupt change in the galvanometer current occurs when V_a approaches 10.4 V. The current however, now does not get reversed as before when V is reversed. This can be explained if we take $10.4V$ as an ionisation potential of mercury. Inelastic collisions of the electrons of energy 10.4 eV with the mercury atoms in the collision region produce poitive ions. These positive ions will be accelerated by the large P.D. = V_r, and will reach the plate, whatever be direction of the comparatively small P.D. = V. The current recorded by G is now due to positive ions. This ionisation current cannot be reversed by reversing the direction of V. Hence 4.7 and 6.7 volts are excitation potentials, while 10.4 volts is the ionisation potential of mercury. This method thus enables one to distinguish between excitation and ionisation potentials.

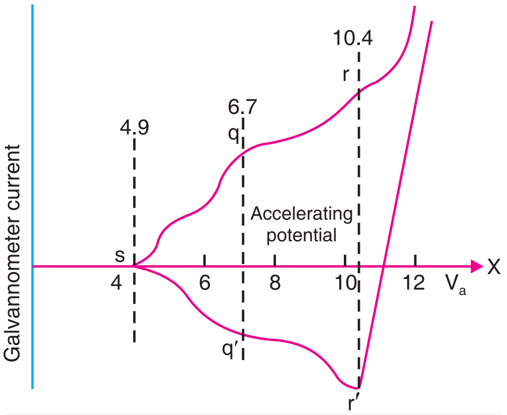


Fig. 6.15

This method is subject to a systematic error on account of the initial velocity of emission of the electrons from the hot filament. The actual energy of the electrons will therefore be greater than that corresponding to the accelerating potential. The correction for this error can be determined by measuring known critical potentials with the apparatus used.

6.11 Sommerfeld's Relativistic Atom Model

Introduction. According to Bohr, the lines in the hydrogen spectrum should each have a well-defined wavelength. Spectrographs of high resolving power showed that the H_α, H_β, and H_γ lines in the hydrogen spectrum are not single. Each spectral line actually consisted of several very close lines packed together. Michelson found that under high resolution, the H_α line can be resolved

Arnold Sommerfeld
(1868 - 1951)

into two close components, with a wavelength separation of 0.13 Å. This is called the *fine structure of the spectral lines*. Bohr's theory could not explain this fine structure.

To explain the observed fine structure of spectral lines, Sommerfeld introduced two main modifications in Bohr's theory.

(1) According to Sommerfeld, the path of an electron around the nucleus, in general, is an ellipse with the nucleus at one of the foci. The circular orbits of Bohr are a special case of this.

(2) The velocity of the electron moving in an elliptical orbit varies considerably at different parts of the orbit. This causes relativistic variation in the mass of the moving electron. Therefore he took into account the relativistic variation of the mass of the electron with velocity. Hence this is known as the relativistic atom-model. Let us first analyse mathematically Sommerfeld's theory.

Elliptic orbits for hydrogen. Consider the electron moving in an elliptical orbit round the nucleus (N). Its position at any instant can be fixed in terms of polar coordinates r and ϕ (Fig. 6.16). In the case of circular orbits, there is only one coordinate that varies periodically, namely, the angle ϕ that the radius vector makes with the X-axis. In the case of elliptic motion, not only does the angle ϕ vary but the length of the radius vector r also varies periodically, as shown in Fig. 6.16. We have now to quantise the momenta associated with both these coordinates (ϕ and r) in accordance with Bohr's quantum condition. The two quantisation conditions are

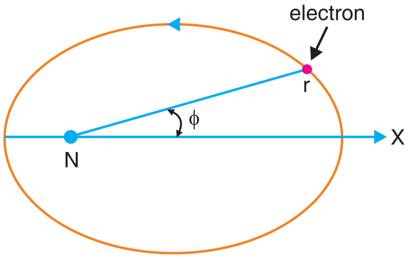


Fig. 6.16

$$\int p_\phi \, d\phi = n_\phi h \qquad \text{...(1)}$$

and
$$\int p_r \, dr = n_r h \qquad \text{...(2)}$$

n_ϕ is called the *angular* or *azimuthal* quantum number and n_r is called the *radial* quantum number. n_ϕ and n_r take only positive integral values and $n_\phi + n_r = n$, where n is the principal quantum number. It can take the integral values 1, 2, 3, ... etc.

To determine the allowed elliptical orbits, we have to evaluate the integrals in equation (1) and (2). The momentum p_ϕ corresponding to the coordinate ϕ is merely the angular momentum p of the electron in the elliptic orbit and this, from Kepler's law, is a constant, *i.e.*, $p_\phi = p =$ constant. Integrating equation (1) from 0 to 2π,

$$\int_0^{2\pi} p_\phi \, d\phi = n_\phi h$$

or
$$p_\phi = \frac{n_\phi h}{2\pi} \qquad \text{...(3)}$$

Evaluation of $\oint p_r \, dr = n_r h$

Now, momentum along the radius $= p_r = m \dfrac{dr}{dt}$

Now
$$p_r \, dr = m \frac{dr}{dt} dr = m \left(\frac{dr}{d\phi} \frac{d\phi}{dt} \right) \frac{dr}{d\phi} d\phi = m \left(\frac{dr}{d\phi} \right)^2 \frac{d\phi}{dt} d\phi$$

But
$$p_\phi = mr^2 \frac{d\phi}{dt}$$

$$\therefore \qquad p_r\,dr = \left(\frac{1}{r}\frac{dr}{d\phi}\right)^2 \; p_\phi\; d\phi \qquad \text{...(4)}$$

The equation of an ellipse in polar coordinates is

$$\frac{1}{r} = \frac{1+\varepsilon\cos\phi}{a\,(1-\varepsilon^2)} \qquad \text{...(5)}$$

where a is the semimajor axis and ε is the eccentricity.

Taking logarithm of Eqn. (5) and differentiating,

$$\frac{1}{r}\frac{dr}{d\phi} = \frac{\varepsilon\sin\phi}{1+\varepsilon\cos\phi} \qquad \text{...(6)}$$

$$\therefore \qquad p_r\,dr = \frac{\varepsilon^2\sin^2\phi}{(1+\varepsilon\cos\phi)^2}\, p_\phi\, d\phi \qquad \text{...(7)}$$

Eqn. (2) now becomes $p_\phi \int_0^{2\pi} \frac{\varepsilon^2\sin^2\phi}{(1+\varepsilon\cos\phi)^2}\,d\phi = n_r h$ ...(8)

The integral $I = \int_0^{2\pi} \frac{\varepsilon^2\sin^2\phi}{(1+\varepsilon\cos\phi)^2}\,d\phi$ can be integrated by parts.

$$\int u\,dv = uv - \int v\,du\,.$$

Let $u = \varepsilon\;\sin\phi$, $du = \varepsilon\;\cos\phi\;d\phi$ and

$$dv = \frac{\varepsilon\sin\phi}{(1+\varepsilon\cos\phi)^2}\,d\phi \qquad \text{so that} \quad v = \frac{1}{1+\varepsilon\cos\phi}$$

Then

$$I = \left[\frac{\varepsilon\sin\phi}{1+\varepsilon\cos\phi}\right]_0^{2\pi} - \int_0^{2\pi}\frac{\varepsilon\cos\phi}{1+\varepsilon\cos\phi}\,d\phi\,.$$

On substitution of the limits of integration, the first term on the R.H.S. becomes zero. Then

$$I = -\int \frac{\varepsilon\cos\phi}{1+\varepsilon\cos\phi}\quad \phi$$

$$= \int_0^{2\pi}\left(\frac{1}{1+\varepsilon\cos\phi} - 1\right)d\phi$$

Now $\int_0^{2\pi}\frac{d\phi}{1+\varepsilon\cos\phi}$ is a standard integral whose value works out to be $2\pi / (1-\varepsilon^2)^{1/2}$.

$$\therefore \qquad I = \frac{2\pi}{(1-\varepsilon^2)^{1/2}} - 2\pi \qquad \text{...(9)}$$

Now, Eqn. (8) can be written as

$$\frac{2\pi\,p_\phi}{(1-\varepsilon^2)^{1/2}} - 2\pi\,p_\phi = n_r\,h$$

$$\text{or} \qquad \frac{n_\phi h}{(1-\varepsilon^2)^{1/2}} - n_\phi h = n_r\,h\left(\text{since } p_\phi = n_\phi\,\frac{h}{2\pi}\right). \qquad \text{...(10)}$$

$$\therefore \qquad n_r = \frac{n_\phi}{(1-\varepsilon^2)^{1/2}} - n_\phi \quad \text{or} \quad n_r + n_\phi = \frac{n_\phi}{(1-\varepsilon^2)^{1/2}}$$

But, $n_r + n_\phi = n$, the principal quantum number.

Hence,
$$n = \frac{n_\phi}{(1-\varepsilon^2)^{1/2}} \quad \text{or} \quad 1-\varepsilon^2 = \frac{n_\phi^2}{n^2} \qquad \text{...(11)}$$

For an ellipse $1 - \varepsilon^2 = b^2/a^2$ where a and b are the semi-major and semi-minor axes respectively.

Hence,
$$\frac{b^2}{a^2} = \frac{n_\phi^2}{n^2}$$

or
$$\frac{b}{a} = \frac{n_\phi}{n} \qquad \text{...(12)}$$

Equation (11) is the condition that determines the allowed elliptical orbits. The allowed elliptical orbits are those for which the ratio of major and minor axes is that of two integers. When $n_\phi = n$, $b = a$, $\varepsilon = 0$ and the orbit becomes circular. n_ϕ cannot be zero, since the ellipse would then degenerate into a straight line passing through the nucleus. Also n_ϕ cannot be greater than n, since b is always less than a. Hence for a given value of n, n_ϕ can assume only n different values, which means there can be only n elliptical orbits of different eccentricities.

Example. *For the first orbit, $n = 1$. Since $n_r + n_\phi = 1$ and $n_\phi \neq 0$, $n_\phi = 1$. With $n = n_\phi = 1$, the first orbit is a circle. With $n = 2$, n_ϕ may have the values 1 or 2, so that there are two possible orbits for $n = 2$, a circle and an ellipse. Similarly, there are three possible orbits for $n = 3$, a circle and two ellipses* (*Fig. 6.17*).

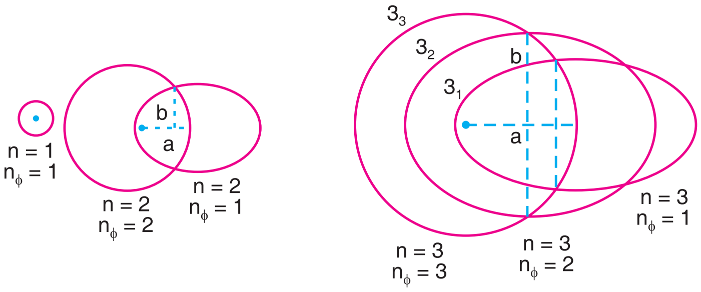


Fig. 6.17

Usually the allowed orbits are described by giving values of n and n_ϕ. The three orbits for $n = 3$ are represented by symbols 3_3, 3_2 and 3_1, the subscript being the azimuthal quantum number (n_ϕ).

Another notation. In this notation, the value of azimuthal quantum number n_ϕ is described by the letters, s, p, d, f, etc. The value of n_ϕ corresponding to these letters is 1, 2, 3, 4 etc., respectively. In this notation, the orbit determined by $n = 3$ and $n_\phi = 1$ is represented by 3s. Similarly 4d will represent the orbit $n = 4$ and $n_\phi = 3$.

Total energy. The total energy E_n of a single electron = E_n = *P.E.* + *K.E.*

The *P.E.* is $\dfrac{-Ze^2}{4\pi\varepsilon_0 r}$.

The *K.E.* can be written as $\dfrac{1}{2}m\left[\left(\dfrac{dr}{dt}\right)^2 + \left(r\dfrac{d\phi}{dt}\right)^2\right]$

where $\dfrac{dr}{dt}$ is the radial component of the velocity and $r\left(\dfrac{d\phi}{dt}\right)$ is the transverse component of the velocity. Hence

$$E_n = \frac{-Ze^2}{4\pi\varepsilon_0 r} + \frac{1}{2} m \left[\left(\frac{dr}{dt}\right)^2 + \left(r\frac{d\phi}{dt}\right)^2 \right] \qquad ...(13)$$

It can be shown that

$$E_n = -\left(\frac{me^4Z^2}{8\varepsilon_0^2 h^2}\right)\frac{1}{n^2} \qquad ...(14)$$

We find that the expression for the total energy is the same as that obtained by Bohr. This means that the theory of elliptical orbits introduces no new energy levels, other than those given by Bohr's theory of circular orbits. No new spectral lines, which would explain the fine structure, are to be expected because of this *multiplicity* of orbits. Hence, Sommerfeld proceeded further to find a solution to the problem of fine structure of spectral lines, on the basis of the variation of the mass of the electron with velocity.

Sommerfeld's relativistic theory. The velocity of the electron in a circular orbit is constant. But the velocity of the electron in an elliptical orbit varies, being a maximum at the perihelion and a minimum at the aphelion. Furthermore this velocity is quite large $\left(\approx \frac{1}{137}c\right)$. Therefore, Sommerfeld modified his theory, taking into account variation of the mass of the electron with velocity. He showed that the relativistic equation describing the path of an electron is

$$\frac{1}{r} = \frac{1+\varepsilon \cos \psi\phi}{a(1-\varepsilon^2)} \qquad ...(15)$$

where $\psi^2 = 1 - \frac{Z^2e^4}{16\pi^2 \varepsilon_0^2 p^2 c^2}$.

The path of the electron given by equation (15) is a *rosette* — a *precessing ellipse. i.e.*, an ellipse whose major axis precesses slowly in the plane of the ellipse about an axis through the nucleus (Fig. 6.18).

It can be shown that the total energy E_n in the relativistic theory is,

$$E_n = \frac{me^4Z^2}{8\varepsilon_0^2 h^2 n^2} - \frac{me^4Z^4\alpha^2}{8\varepsilon_0^2 h^2}\left(\frac{n}{n_\phi} - \frac{3}{4}\right)\frac{1}{n^4} \qquad ...(16)$$

where $\alpha = \frac{e^2}{2\varepsilon_0 ch} \approx \frac{1}{137}$.

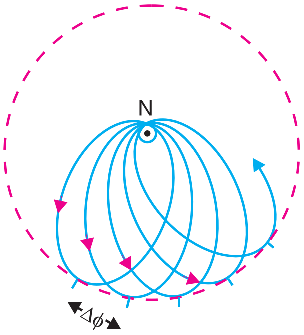


Fig. 6.18

α is a dimensionless quantity and is called the *fine structure constant.* The first term in equation (16) is the energy of the electron in the orbit with the principal quantum number n. The second term is the *Sommerfeld relativity correction.* This term shows that the energy does depend on the azimuthal quantum number n_ϕ. This results in a splitting of the energy levels of the atom, for a given value of n, into n components, corresponding to the n permitted values of n_ϕ.

Fine structure of the H_α line.

H_α line is due to the transition from $n = 3$ state to $n = 2$ state of the hydrogen atom.

For $n = 3$, there are three possible energy levels corresponding to the three values of n_ϕ, 1, 2, and 3. Similarly there are two possible energy levels for $n = 2$. Theoretically, six transitions are possible : $3_3 \rightarrow 2_2$; $3_3 \rightarrow 2_1$; $3_2 \rightarrow 2_2$; $3_2 \rightarrow 2_1$; $3_1 \rightarrow 2_2$; $3_1 \rightarrow 2_1$, and they are shown in Fig. 6.19.

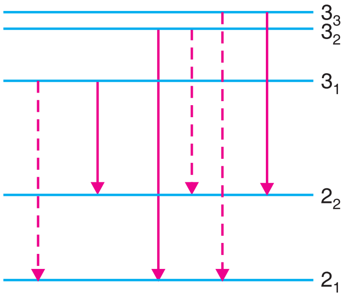


Fig. 6.19

Actually the H_α line has only *three* components. To make experiment and theory agree, some of the transitions have to be ruled out by some *selection rule.* The selection rule is that n_ϕ can change only by +1 or –1, *i.e.*, $\Delta n_\phi = \pm 1$.

There is no such restriction on Δn. On the basis of this selection rule, there are only three allowed transitions. In Fig. 6.19, the allowed transitions are shown by continuous lines and the forbidden lines by dotted lines.

6.12 The Vector Atom Model

Introduction. (1) Bohr's theory was able to explain only the series spectra of the simplest hydrogen atom. It could not explain the multiple structure of spectral lines in the simplest hydrogen atom. Sommerfeld's theory was able to give an explanation of the fine structure of the spectral lines of hydrogen. However, Sommerfeld's theory could not predict the correct number of the fine structure lines. Moreover, it gave no information about the relative intensities of the lines. Sommerfeld's theory could not explain the complex spectra of alkali metals like sodium.

(2) These older theories were inadequate to explain new discoveries like Zeeman Effect and Stark Effect in which the spectral lines could be split up under the influence of magnetic and electric fields.

(3) Another drawback of the Bohr model was that it could not explain how the orbital electrons in an atom were distributed around the nucleus.

Therefore, in order to explain the complex spectra of atoms and their relation to atomic structure, the vector atom model was introduced. The two distinct features of the vector atom model are:

1. *the conception of spatial quantisation*, and
2. *the spinning electron hypothesis*.

(1) **Spatial quantisation.** According to Bohr's theory, the orbits are quantised as regards their *magnitude* (*i.e.*, their size and form) only. But according to quantum theory, the *direction* or orientation of the orbits in space also should be quantised. To specify the orientation of the electron orbit in space, we need a fixed reference axis. This reference line is chosen as the direction of an external magnetic field that is applied to the atom. The different permitted orientations of an electron orbit are determined by the fact that *the projections of the quantised orbits on the field direction must themselves be quantised.* The idea of space quantisation leads to an explanation of Zeeman effect. The Stern-Garlach experiment provided an excellent proof of the space quantisation of atom.

(2) **Spinning electron.** To account for the observed fine structure of spectral lines and to explain the anomalous Zeeman effect, the concept of *spinning electron* was introduced by Uhlenbeck and Goudsmit in 1926. According to their hypothesis, the electron spins about an axis of its own, while it also moves round the nucleus of the atom in its orbit. The spin of the electron is analogous to the spinning of a planet about its own axis, as it moves in an elliptical orbit around the sun. In other words, the electron is endowed with a spin motion over and above the orbital motion. According to the quantum theory, the spin of the electron also should be quantised. Hence a new quantum number called the *spin quantum number* (s) is introduced. Since the orbital and spin motions are both quantised in magnitude and direction according to the idea of spatial quantisation, they are considered as *quantised vectors.* Hence the atom model based on these *quantised vectors* is called the "vector atom model", to which vector laws apply.

According to the older theories, the electron was supposed to have only orbital motion round the nucleus. Hence, only the orbital angular momentum and orbital magnetic moment were considered. The spin endows the electron with a spin angular momentum $s\hbar$ and a spin magnetic moment. Hence the total angular momentum of an atom should be the vector sum of the orbital angular momentum and spin angular momentum. Similarly, the total magnetic moment of an atom should be the vector sum of the orbital and spin magnetic moments.

6.13 Quantum Numbers Associated with the Vector Atom Model

(1) **The principal quantum number (n).** This is identical with the one used in Bohr-Sommerfeld theory. The serial number of the shells starting from the innermost is designated as its principal quantum number (n). It can take only integral values excluding zero. *i.e.*, $n = 1, 2, 3, 4, \ldots$

(2) **The orbital quantum number (l).** This may take any integral value 0, 1, 2, 3, 4, ..., $(n - 1)$. Thus, if $n = 4$, l can take the four values 0, 1, 2, 3. By convention, an electron for which $l = 0$, is called s electron; if $l = 1$, p electron; $l = 2$, d electron; $l = 3$, f electron etc. The orbital angular momentum p_l of the electron is given by $p_l = l\hbar$.

According to the wave mechanics, $p_l = [l(l + 1)]^{1/2}\hbar$

(3) **The spin quantum number (s).** This has only one values, $= \frac{1}{2}$. The spin angular momentum $p_s = s\hbar$ where $s = \frac{1}{2}$. According to wave mechanics, $p_s = [s\,(s + 1)^{1/2}]\hbar$

(4) **Total angular momentum quantum number (j).** It represents the total angular momentum of the electron which is the sum of the orbital angular momentum and spin angular momentum. The vector $\vec{j}$ is defined by the equation $\vec{j} = \vec{l} + \vec{s}$ with the restriction that $\vec{j}$ is positive. The spin angular momentum $s = \pm\frac{1}{2}$. $\therefore j = l \pm s$, plus sign when s is *parallel* to l and minus sign when s is *antiparallel*. Thus for $l = 2$ and $s = \frac{1}{2}$, j can have the values $\frac{5}{2}$ and $\frac{3}{2}$ (Fig. 6.20).

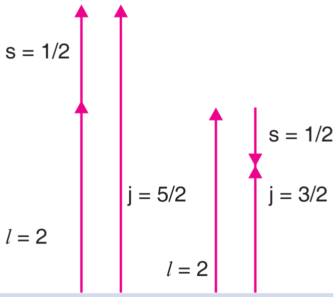


Fig. 6.20

The total angular momentum of the electron $= p_j = j\hbar$.

According to wave mechanics, $p_j = \sqrt{j\,(j+1)}\;\hbar$.

To explain the splitting of spectral lines in a magnetic field, three more magnetic quantum numbers are introduced.

(5) **Magnetic orbital quantum number (m_l).** The projection of the orbital quantum number l on the magnetic field direction is called the magnetic orbital quantum number m_l. The possible values of m_l are $l, l - 1, l - 2, \ldots, 0, -1, -2, \ldots -l$. *i.e.*, there are $(2l + 1)$ possible values of m_l. This is illustrated in Fig. 6.21 for $l = 3$. The angle θ between l and B is given by $\cos\,\theta = \frac{m_l}{l}$. Conversely, the permitted orientations of the l vector relative to the field direction B is also $(2l + 1)$. For example, if $l = 3$, the permitted orientations of l are 7, for which $m_l = 3, 2, 1, 0, -1, -2, -3$. Hence, the l vector can take only the 7 directions shown in Fig. 6.21. l cannot be inclined to B at any other angle. This is known as '*spatial quantisation*'.

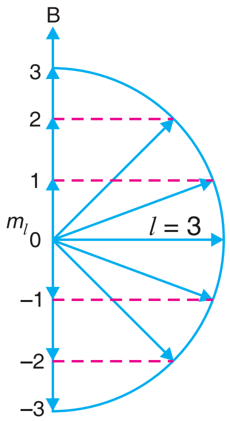


Fig. 6.21

(6) **Magnetic spin quantum number (m_s).** This is the projection of the spin vector s along the direction of the magnetic field. The spin angular momentum (s) can assume only two possible positions with respect to the magnetic field: it may be parallel to it or antiparallel. m_s can have only two values $+\frac{1}{2}$ or $-\frac{1}{2}$, as illustrated in Fig. 6.22.

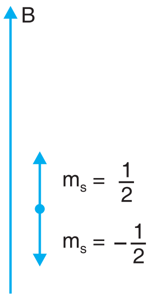


Fig. 6.22

(7) **Magnetic total angular momentum quantum number (m_j).** This is the projection of total angular momentum vector j on the direction of the magnetic field. Since we are dealing with a single electron, j can have only odd half-integral values $\left(\because \quad j = l \pm \frac{1}{2}\right)$. Hence, m_j must have only odd half-integral values. m_j can have only $(2j + 1)$ values, from $+j$ to $-j$ zero excluded. Fig. 6.23 shows the possible values of m_j for $j = 3/2$.

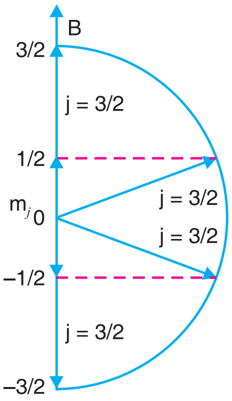


Fig. 6.23

Notes

(1) The state of an electron in an atom is completely specified by the four quantum numbers, n, l, m_l, and m_s.

(2) In spectroscopic notation, small letters, l, s, j and s, p, d, f, g etc., are used to describe the state of the electron and capital letters, L, S, J, and S, P, D, F, G, etc., are used to describe the state of the atom as a whole.

6.14 Coupling Schemes

In an atom having two or more electrons, the orbital and spin angular momenta of all its electrons can be added together in two ways. The method of combination depends on the interaction or *coupling* between the orbital and spin angular momenta. Two types of schemes have been developed. They are: (1) L—S coupling or the Russel-Saunders coupling and (2) the j—j coupling.

(1) **L–S coupling.** The type of coupling which occurs most frequently is the L—S coupling. In this type, all the orbital angular momentum vectors of the various electrons combine to form a resultant L and independently, all their spin angular momentum vectors combine to form a resultant S. These resultants L and S then combine to form the total angular momentum J of the atom (Fig. 6.24). This scheme may be summarised as follows: $L = \sum l_i$; $S = \sum s_i$; $J = L + S$. L is always an integer including zero. S is an integer for an even number of electrons, and odd multiple of $\frac{1}{2}$ for an odd number of electrons (Fig. 6.25). Hence J must be an integer, if S is an integer and J must be an odd multiple of $\frac{1}{2}$ if S is an odd multiple of $\frac{1}{2}$. It can be shown that, when $L > S$, J can have $(2S + 1)$ values and when $L < S$, J can have $(2L + 1)$ values. In particular, if $L = 0$, J can have only one value namely $J = S$.

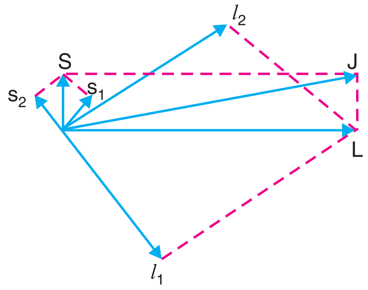


Fig. 6.24

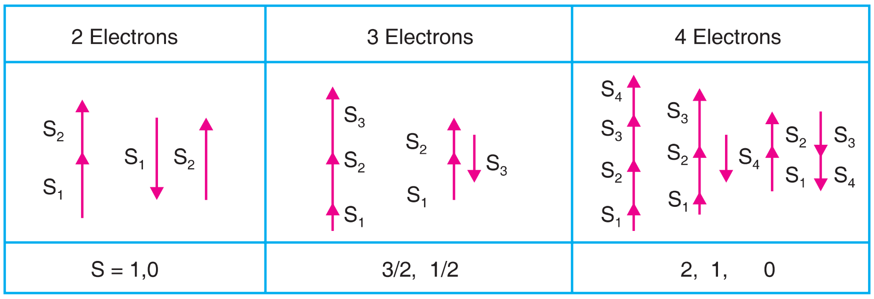


Fig. 6.25

(2) **The j-j coupling.** This method is employed when the interaction between the spin and orbital vectors of each electron is stronger than the interaction between either the spin vectors or the orbital vectors of the different electrons. The orbital and spin angular momenta of each electron in the atom are added to obtain the resultant angular momentum of the electron (Fig. 6.26).

Thus $\vec{j_i} = \vec{l_i} + \vec{s_i}$.

The vector sum of all the individual j vectors gives the total angular momentum J of the atom. Thus $J = \sum j_i$. This type of coupling exists mainly in heavy atoms.

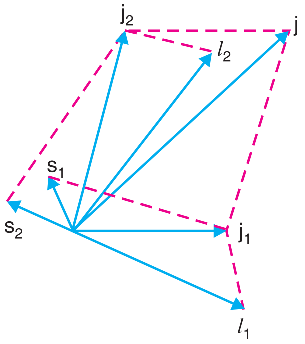


Fig. 6.26

Application of spatial quantisation. The resultant vectors **L, S** and **J,** representing the atom, can be obtained by the above coupling schemes. According to quantum theory **L, S** and **J** are quantised in magnitude and direction. Hence the number of permitted orientations of **L, S** and **J** with respect to a given field direction are $(2L + 1)$, $(2S + 1)$ and $(2J + 1)$ respectively. The corresponding magnetic quantum numbers $m_L = \sum m_l$, $m_S = \sum m_s$ and $m_J = \sum m_j$, can have only $(2L + 1)$, $(2S + 1)$ and $(2J + 1)$ values respectively.

Note

For a one-electron atom with only a single effective electron, the state of the atom as a whole is identical with the state of the electron, *i.e.,* $L = l$. $S = s$ and $J = j$.

6.15 The Pauli Exclusion Principle

Statement. *No two electrons in an atom exist in the same quantum state.* The four quantum numbers n, l, m_l and m_s determine the state of an electron completely. Hence the principle may be stated as "*No two electrons in an isolated atom may have the same four quantum numbers.*"

Explanation. The principle implies that each electron in an atom must have a different set of quantum numbers n, l, m_l and m_s. If two electrons have all their quantum numbers identical, then one of those two electrons would be excluded from entering into the constitution of the atom. Hence the name "*exclusion principle*".

Application. This principle enables us to calculate the number of electrons that can occupy a given subshell.

Wolfgang Pauli (1900 – 1958)

(1) Consider the K-shell with $n = 1$, $l = 0$ and hence $m_l = 0$. Since $s = \frac{1}{2}$, m_s can be either $+\frac{1}{2}$ or $-\frac{1}{2}$. Hence, the K-shell can have two electrons: *electron* 1 *with quantum numbers* $n = 1$, $l = 0$,

$m_l = 0,\ m_s = \frac{1}{2}$; and electron 2 with quantum numbers $n = 1, l = 0, m_l = 0,\ m_s = -\frac{1}{2}$. If there were a third electron, its quantum numbers will be identical with those of the first or second electron, which is against Pauli's exclusion principle. The *K*-shell is therefore completed or closed with two electrons.

(2) For the *L*-shell, $n = 2$ and $l = 0$ or 1. For the subshell $n = 2,\ l = 0,\ m_l$ must be zero and $m_s = +\frac{1}{2}$ or $-\frac{1}{2}$. Hence there can be only 2 electrons in this subshell. For the subshell $n = 2$, $l = 1$, m_l can have three values –1, 0, +1. For each of these three values of m_l, m_s may be either $+\frac{1}{2}$ or $-\frac{1}{2}$. Hence there will be six possible set of values for the quantum numbers characterizing the electrons. Therefore, the maximum number of electrons in this subshell is 6. The *L*-shell with two subshells [($n = 2,\ l = 0$) and ($n = 2,\ l = 1$)] is, therefore, completed when it contains 2 + 6 = 8 electrons.

(3) For the *M*–shell with $n = 3$, there can be three subshells with $l = 0, 1, 2$. The first and second subshells are completed by 2 and 6 electrons as explained above. The third subshell is completed with $2(2l + 1)$, *i.e.*, 10 electrons since $l = 2$. Hence the total number of electrons required to complete the *M*-shell is 18.

(4) Similarly, the *N*–shell can have a maximum of 32 electrons.

In general, we get the two following conclusions:

(*i*) In the n^{th} shell there are *n* sub-shells corresponding to the values 0, 1, 2, 3 ... $(n - 1)$ of *l*.

The maximum number of electrons in a sub-shell with a given value of *l* is $2(2l + 1)$.

Orbital quantum number (*l*)	0	1	2	3	4	.....
Number of possible electron states	2	6	10	14	18	.....
Subshell symbol	*s*	*p*	*d*	*f*	*g*	.....

(*ii*) The number of electrons that can be accommodated in a shell with principal quantum number *n* = Sum of the electrons in the constituent *n* subshells.

$$= \sum_{l=0}^{l=n-1} 2(2l+1) = 2\sum_{l=0}^{l=n-1} (2l+1)$$

$$= 2\,[\,1 + 3 + 5 + 7 + \ldots\ldots \{2(n-1) + 1\}] = 2n^2$$

The following table shows the distribution of electrons according to this scheme.

Shell symbol	*K*	*L*	*M*	*N*	*O*
Quantum number (*n*)	1	2	3	4	5
Number of electrons ($2n^2$)	2	8	18	32	50

The distribution of electrons in the various states (shells and sub-shells) according to the exclusion principle is given in the following table (Table 6.1).

TABLE 6.1

n	l	m_l	m_s	*Number of electrons in sub-shell with spectroscopic notation*		*Total number of electrons in shell = $2n^2$*
1	0	0	$+\frac{1}{2}, -\frac{1}{2}$	2	$1s^2$	2
2	0	0	$+\frac{1}{2}, -\frac{1}{2}$	2	$2s^2$	8
2	1	–1, 0, +1	$+\frac{1}{2}, -\frac{1}{2}$	6	$2p^6$	
3	0	0	$+\frac{1}{2}, -\frac{1}{2}$	2	$3s^2$	18
3	1	–1, 0, +1	$+\frac{1}{2}, -\frac{1}{2}$	6	$3p^6$	
3	2	–2, –1, 0, +1, +2	$+\frac{1}{2}, -\frac{1}{2}$	10	$3d^{10}$	
4	0	0	$+\frac{1}{2}, -\frac{1}{2}$	2	$4s^2$	32
4	1	–1, 0, +1	$+\frac{1}{2}, -\frac{1}{2}$	6	$4p^6$	
4	2	–2, –1, 0, +1, +2	$+\frac{1}{2}, -\frac{1}{2}$	10	$4d^{10}$	
4	3	–3,–2,–1,0,+1,+2,+3	$+\frac{1}{2}, -\frac{1}{2}$	14	$4f^{14}$	

6.16 The Periodic Classification of Elements

The periodic table. The periodic table is an arrangement of different elements that exist in nature, based on their chemical properties and atomic numbers. Table 6.2 shows the simplest form of the periodic table.

Elements with similar properties form the *groups* shown as vertical columns in the table. Thus group I consists of hydrogen plus the alkali metals, all of which are extremely active chemically and all of which have valence of +1. Group VII consists of the halogens that have valence of –1. Group VIII consists of the inert gases which are chemically inactive.

The horizontal rows are called *periods*. As we go from left to right in the same period, the chemical and physical properties of the elements vary gradually as the atomic number increases. Since the atomic number gives also the number of electrons in the atom, it follows that the atoms of successive elements in the periodic table are formed by the addition of one more electron at each step.

We have already seen the arrangement of electrons in an atom by applying Pauli's exclusion principle. The notion of electron shells and subshells fits perfectly into the pattern of the periodic table. The total orbital and spin angular momenta of the electrons in a closed subshell are zero. The electrons in a closed shell are all very tightly bound, since the positive nuclear charge is large relative to the negative charge of the inner shielding electrons. Since an atom containing only closed shells has no dipole moment, it does not attract other electrons, and its electrons, cannot be readily detached. We expect such atoms to be passive chemically, like the inert gases and the inert gases all turn out to have closed shell electron configurations.

TABLE 6.2

Period	*Group I*	*Group II*											*Group III*	*Group IV*	*Group V*	*Group VI*	*Group VII*	*Group VIII*
1	1 **H** 1.00																	2 **He** 4.00
2	3 **Li** 6.94	4 **Be** 9.01											5 **B** 10.81	6 **C** 12.01	7 **N** 14.01	8 **O** 16.00	9 **F** 19.00	10 **Ne** 20.18
3	11 **Na** 22.99	12 **Mg** 24.31											13 **Al** 26.48	14 **Si** 28.09	15 **P** 30.98	16 **S** 32.07	17 **Cl** 35.46	18 **Ar** 39.94
4	19 **K** 39.10	20 **Ca** 40.08	21 **Sc** 44.96	22 **Ti** 47.90	23 **V** 50.94	24 **Cr** 52.00	25 **Mn** 54.94	26 **Fe** 55.85	27 **Co** 58.93	28 **Ni** 58.71	29 **Cu** 63.54	30 **Zn** 65.37	31 **Ga** 69.72	32 **Ge** 72.59	33 **As** 74.92	34 **Se** 78.96	35 **Br** 79.91	36 **Kr** 83.98
5	37 **Rb** 85.47	38 **Sr** 87.66	39 **Y** 88.91	40 **Zr** 91.22	41 **Nb** 92.91	42 **Mo** 95.94	43 **Tc** (99)	44 **Ru** 101.1	45 **Pd** 102.91	46 **Ag** 106.4	47 **Cd** 107.87	48 **In** 112.82	49 **Sn** 114.82	50 **Sb** 118.69	51 **Te** 121.75	52 **I** 127.60	53 **Xe** 126.90	54 131.30
6	55 **Cs** 132.91	56 **Ba** 137.34	57-71 *	72 **Hf** 178.49	73 **Ta** 180.95	74 **W** 183.85	75 **Re** 186.2	76 190.0	77 **Ir** 192.2	78 **Pt** 195.09	79 **Au** 197.0	80 **Hg** 200.59	81 **Tl** 204.37	82 **Pb** 207.19	83 **Bi** 208.98	84 **Po** (210)	85 **At** (210)	86 **Rn** 222
7	87 **Fr** (223)	88 **Ra** 226.05	89-103 **															
* Rare earths			57 **La** 138.91	58 **Ce** 140.12	59 **Pr** 140.91	60 **Nd** 144.24	61 **Pm** (145)	62 **Sm** 150.35	63 **Eu** 152.0	64 **Gd** 157.25	65 **Tb** 158.92	66 **Dy** 162.50	67 **Ho** 164.92	68 **Er** 167.26	69 **Tm** 168.93	70 **Yb** 173.04	71 **Lu** 174.97	
** Actinides			89 **Ac** (227)	90 **Th** 232.04	91 **Pa** (231)	92 **U** 238.03	93 **Np** (237)	94 **Pu** (242)	95 **Am** (243)	96 **Cm** (247)	97 **Bk** (249)	98 **Cf** (251)	99 **Es** (254)	100 **Fm** (253)	101 **Md** (256)	102 **No** (254)	103 **LW** (257)	

Those atoms which have a single electron in their outermost shell, tend to lose this electron. Hydrogen and the alkali metals are in this category and accordingly have valences of +1. Atoms whose outer shells lack a single electron for being closed, tend to acquire such an electron, which accounts for the chemical behaviour of the halogens. It is clear that the chemical and physical properties of an atom are determined by the number and arrangement of the electrons in the *outermost-shell* and not by the total number of electrons in the atom. In this manner the similarities of the members of the various groups of the periodic table may be accounted for.

6.17 Some Examples of Electron Configurations with their Modern Symbolic Representations

The *electronic configuration* of an atom is the distribution of electrons in various subshells around the nucleus of the atom. In describing the electron configuration, small letters are used to represent the values of l as follows:

$$\begin{array}{lcccccc} l = & 0, & 1, & 2, & 3, & 4, & 5, \ldots\ldots \\ & s, & p, & d, & f, & g, & h, \ldots\ldots \end{array}$$

i.e., if an electron is in a shell for which $l = 0$, it is called an s electron; for $l = 1$, a p electron and so on. The value of the principal quantum number n is written as a prefix to the letter representing its l value. *For example*, a state in which $n = 2$, $l = 0$ is a $2s$ – state; $n = 4$, $l = 2$ is a $4d$–state and so on. The number of electrons having the same n and l values is indicated by an index written at the upper right of the letter representing their l value. Thus the 11 electrons of sodium in the normal state are designated as follows: $1s^2\ 2s^2\ 2p^6\ 3s$. *i.e.*, there are two $1s$ electrons, two $2s$ electrons, six $2p$ electrons, and one $3s$ electron. We shall now consider electron configurations of a few elements.

(1) **Hydrogen (Z = 1).** The normal state of an atom is one in which all the electrons are in the lowest possible energy levels. In hydrogen, the normal state is characterised by the quantum numbers $n = 1$, $l = 0$, $m_l = 0$ and m_s may be either $+\frac{1}{2}$ or $-\frac{1}{2}$. The symbolic representation is $1s$. The K–shell requires one more electron to be completed. Hence atomic hydrogen is very active chemically.

(2) **Helium (Z = 2).** It has both its electrons in the shell $n = 1$, $l = 0$. $m_s = \frac{1}{2}$ for one electron and $-\frac{1}{2}$ for the second electron. The symbolic representation is $\boxed{1s^2}$. This shell is now completed or closed. The rectangular enclosure indicates that the electrons are interlocked in a closed shell. Therefore, helium may be expected to have a very stable configuration. This should also be true of all the other inert gases.

(3) **Lithium (Z = 3).** It has three electrons. Two electrons can be put in the shell $n = 1$, $l = 0$. The third electron must be put into a new shell $n = 2$, $l = 0$. So the neutral lithium atom is represented by $1s^2\ 2s$. Lithium is one of the alkali elements and has a valence of unity. This means that the valence $2s$ electron can be detached easily from the atom to form the lithium ion Li^+. This is indicated by the fact that its ionisation potential is only 5.39 volts, whereas for *He* it is 24.58 volts. Lithium is chemically quite active and is monovalent. Similarly, all alkali metals (*Na*, *K*, *Rb*, *Cs*) have one electron in their outermost shell and hence are monovalent having similar chemical properties.

(4) **Beryllium (Z = 4).** It has two electrons in the completed K–shell ($n = 1$). It has two additional electrons in the ($n = 2$, $l = 0$) subshell. It is represented by $1s^2\ 2s^2$. Beryllium is one of the alkaline earth elements with a valence of 2. Its optical spectrum is that of a two-electron atom.

The atoms of the other elements of the group (*Mg*, *Ca*, *Sr*, *Ba*, *Ra*), should have similar structures. They have two electrons outside an inert gas or closed shell configuration. They have very similar chemical and physical properties.

Similarly the electronic configurations from boron (Z = 5) to neon (Z = 10) are:

(*i*) Boron : $1s^2\ 2s^2\ 2p$

(*ii*) Carbon (Z = 6) : $1s^2\ 2s^2\ 2p^2$

(*iii*) Nitrogen (Z = 7) : $1s^2\ 2s^2\ 2p^3$

(*iv*) Oxygen (Z = 8) : $1s^2\ 2s^2\ 2p^4$

(*v*) Flourine (Z = 9) : $1s^2\ 2s^2\ 2p^5$

(*vi*) Neon (Z = 10) : $1s^2\ 2s^2\ 2p^6$

In neon, both the $n = 1$ and $n = 2$ shells are completed. Neon is one of the inert gases and has a very stable configuration.

The next eight elements from sodium (Z = 11) to argon (Z = 18) are formed by adding the additional electrons to the *M* shell for which $n = 3$.

(*i*) Sodium (Z = 11) : $\boxed{1s^2\ 2s^2\ 2p^6}\ 3s$. Sodium has an electron ($3s$ electron) outside a closed shell. This single electron, like that in lithium, is easily ionised; the valence is 1; the spectrum is that of one-electron atom.

(*ii*) Magnesium (Z = 12) : $\boxed{1s^2\ 2s^2\ 2p^6}\ 3s^2$. The two electrons in the outermost incomplete *M* shell ($n = 3$) are the valence electrons making *Mg divalent.*

(*iii*) Aluminium (Z = 13) : $\boxed{1s^2\ 2s^2\ 2p^6}\ 3s^2\ 3p$. *Al* is trivalent.

6.18 Magnetic Dipole Moment Due to Orbital Motion of the Electron

Consider an electron moving in an elliptical orbit of area *A* with a period *T* (Fig. 6.27). The electron crosses any point in the orbit $1/T$ times in unit time. This is equivalent to a current $i = e/T$ in a loop of area *A*, where *e* is charge of the electron. Applying Ampere's theorem, this current gives rise to a magnetic dipole moment μ_l given by

$$\mu_l = iA = eA/T \qquad \text{...(1)}$$

where *A* is the area of the orbit.

Since the areal velocity in a central orbit is $\frac{1}{2} r^2 \frac{d\phi}{dt}$, the area

$$A = \int_0^T \frac{1}{2} r^2 \left(\frac{d\phi}{dt}\right) dt.$$

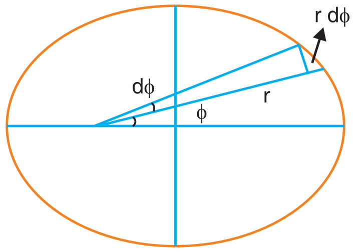


Fig. 6.27

Now, the angular momentum of the electron

$$p_l = mr^2 \frac{d\phi}{dt} = constant$$

or

$$\frac{1}{2} r^2 \frac{d\phi}{dt} = \frac{p_l}{2m} = constant$$

$$\therefore \qquad A = \int_0^T \left(\frac{p_l}{2m}\right) dt = \frac{p_l T}{2m} \qquad \text{...(2)}$$

Substituting this value of A in (1), $\mu_l = \frac{e}{2m} p_l$ ...(3)

Since angular momentum is quantised, we have $p_l = l\hbar$

$$\therefore \qquad \mu_l = \frac{e}{2m} \times l\hbar = \frac{e\hbar}{2m} l \qquad ...(4)$$

μ_l is directly proportional to l.

$e\,\hbar/(2m)$ is the smallest unit of magnetic dipole moment and is called the *Bohr electron magneton* (μ_B).

$$\mu_B = \frac{e\hbar}{2m} = \frac{(1.602 \times 10^{-19}) \times (1.055 \times 10^{-34})}{2 \times (9.109 \times 10^{-31})} = 9.274 \times 10^{-24} \text{ JT}^{-1}$$

Since the electronic charge is negative, the magnetic dipole moment vector is directed opposite to that of the angular momentum (Fig. 6.28). The ratio (magnetic dipole moment/angular momentum) is called the *gyromagnetic ratio.*

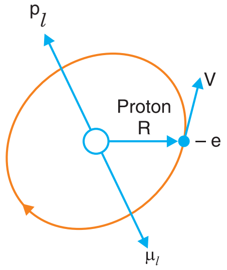


Fig. 6.28

i.e.,

$$\frac{\mu_l}{p_l} = \frac{e}{2m} = 8.8 \times 10^9 \ C/kg. \qquad ...(5)$$

Eq. (5) may now be written as

$$\frac{\mu_l}{p_l} = -g_l \frac{e}{2m} \text{ with } g_l = 1 \qquad ...\ (5a)$$

Similarly,

$$\frac{\mu_s}{S} = -g_s \frac{e}{2m} \text{ with } g_s = 2 \qquad ...(5b)$$

$$\frac{\mu_j}{J} = -g_j \frac{e}{2m} \qquad ...(5c)$$

Here, g is called *Lande's splitting factor.*

Example 1. *A beam of electrons enters a uniform magnetic field of* 1.2 *T. Calculate the energy difference between electrons whose spins are parallel and antiparallel to the field.*

(*Gorakhpur* 87; *Lucknow* 87; *Allahabad* 82).

Sol. An electron has an intrinsic (spin) angular momentum **S** and an intrinsic magnetic dipole moment μ_s, which are related by

$\mu_s = -(e/m)$ S

Let the magnetic field **B** be along the z-axis. The magnitude of the z-component of the magnetic moment is

$$\mu_{sz} = \frac{e}{m} S_z \quad \text{where} \quad S_z = m_s \hbar$$

Here, $m_s = \pm 1/2$, depending upon whether S_z is parallel or antiparallel to the z-axis. Thus

$$\mu_{sz} = \frac{e}{m} m_s \hbar = \frac{e}{m}\left(\pm\frac{1}{2}\right)\hbar = \pm\frac{e\hbar}{2m}.$$

Now, the magnetic potential energy of a dipole of moment μ, in a magnetic field **B** is given by

$$U_m = -\mu_s \cdot \mathbf{B} = -\mu_{sz} \mathbf{B}$$

Here, μ_{sz} is the scalar magnitude of μ_s, in the direction of the magnetic field. Thus

$$U_m = \pm \frac{e\hbar}{2m} B.$$

The difference in energy of the electrons having spin parallel and antiparallel to the field is

$$\Delta U_m = \pm \frac{e\hbar B}{2m} - \left(-\frac{e\hbar B}{2m}\right) = \frac{e\hbar B}{m}$$

$$\therefore \quad \Delta U_m = \frac{(1.6 \times 10^{-19}) \times (1.055 \times 10^{-34}) \times (1.2)}{(9.1 \times 10^{-31})} \text{ J}$$

$$= 2.23 \times 10^{-23}\, J = 1.39 \times 10^{-4} \text{ eV}.$$

Note

When a magnetic dipole μ is placed in an external magnetic field B_{ext}, it experiences a torque (Fig. 6.29).

$$\vec{\tau} = \vec{\mu} \times \vec{B}_{ext}.$$

The torque tends to align the dipole moment vector along that of the field. Maximum torque occurs when the angle θ between μ and B_{ext} is 90°. The P.E. of a magnetic dipole, at any angle θ relative to an external magnetic field is

$$\Delta U_m = -\vec{\mu} \cdot \vec{B}_{ext} = -\mu\, B_{ext} \cos\theta.$$

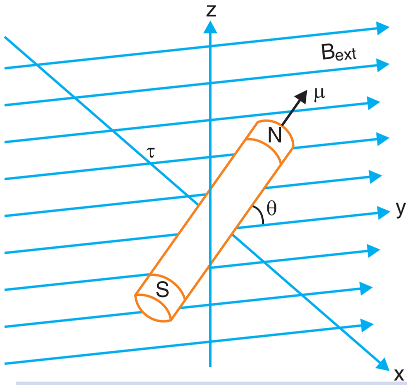


Fig. 6.29

6.19 Magnetic Dipole Moment Due to Spin

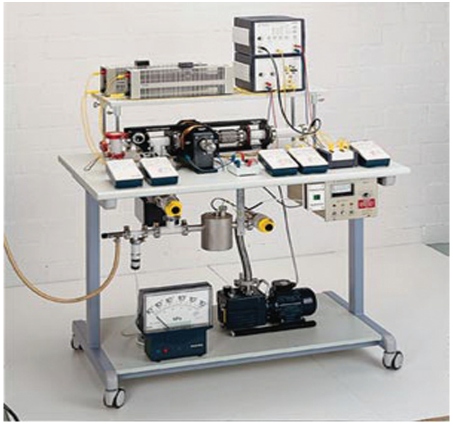
An electron spinning about its axis should also behave as a tiny magnet and possess a magnetic dipole moment due to this spin. However, nothing is known about the shape of an electron or the manner in which its charge is distributed. Hence it is impossible to calculate its spin magnetic dipole moment in a manner analogous to that used for the orbital motion. In order to obtain agreement with experimental results, the spin magnetic dipole moment (μ_s) is assigned the value,

$$\mu_s = 2 \times \frac{e}{2m} p_s \quad \text{where } p_s = s\hbar.$$

$$\therefore \quad \mu_s = 2\frac{e}{2m} \times s\hbar = 2s\frac{e\hbar}{2m} = \frac{e\hbar}{2m} \quad \left(\because s = \frac{1}{2}\right)$$

6.20 The Stern and Gerlach Experiment

We have seen that the orbital and spin motions of the electrons in atoms endow the atoms with magnetic moments. Direct evidence for the existence of magnetic moments of atoms and their space quantisation is provided by the experiments of Stern and Gerlach.



Stern – Gerlach Experiment.

Principle. The experiment is based on the behaviour of a magnetic dipole (atomic magnet) in a *non-uniform magnetic field.* In a uniform magnetic field (*B*), the dipole experiences a torque that tends to align the dipole parallel to the field. If the dipole moves in such a field in a direction normal to the field, it will trace a straight line path without any deviation. In an inhomogeneous magnetic field, the dipole experiences, in addition, a translatory force. If the atomic magnet flies across such an inhomogeneous magnetic field normal to the field direction, it will be deviated away from its rectilinear path. An expression for the deviation produced may be obtained as follows.

Let the magnetic field vary along the X-direction, so that the field gradient is dB/dx and is positive (Fig. 6.30). CD is the atomic magnet (of pole strength p, length l, dipole moment M) with its axis inclined at an angle θ to the field direction. If the field strength at the pole C is B, then the field strength at the other pole D will be $B + \frac{dB}{dx} l \cos\theta$. Hence the forces on the two poles are pB and $p\left(B + \frac{dB}{dx} l \cos\theta\right)$. Hence the atomic magnet experiences not only a torque $(= plB = \mu_s B)$ but also a translatory force

$$F_x = \frac{dB}{dx} pl \cos\theta.$$

$$\therefore \qquad F_x = \frac{dB}{dx} \mu_s \cos\theta \qquad \text{...(1)}$$

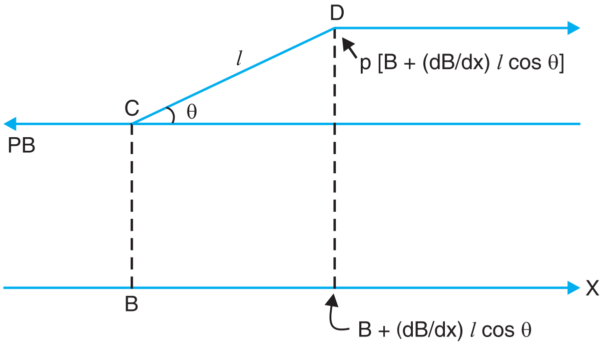


Fig. 6.30

Let V = velocity of the atomic magnet of mass m as it enters the field,
L = length of the path of the atom in the field and
t = the time of travel of the atom through the field = L/V.

$$\left.\begin{array}{l}\text{The acceleration given to}\\ \text{the atom along the field}\\ \text{direction, by the translatory}\\ \text{force } F_x\end{array}\right\} = \frac{F_x}{m} \qquad \text{...(2)}$$

$$\left.\begin{array}{l}\text{The displacement of the}\\ \text{atom along the field}\\ \text{direction, on emerging out}\\ \text{of the field}\end{array}\right\} = \alpha$$

$$= \frac{1}{2}\left(\frac{F_x}{m}\right)t^2 = \frac{1}{2}\frac{F_x}{m}\frac{L^2}{V^2}$$

$$= \frac{1}{2}\frac{dB}{dx}\frac{\mu_s \cos\theta}{m}\frac{L^2}{V^2} \qquad \text{...(3)}$$

If μ is resolved component of the magnetic moment in the field direction, $\mu = \mu_s \cos\theta$

$$\therefore \qquad \alpha = \frac{1}{2}\frac{dB}{dx}\frac{\mu}{m}\frac{L^2}{V^2} \qquad \text{...(4)}$$

Experimental Arrangement. Silver is boiled in an oven [Fig. 6.31 (*a*)]. Atoms of silver stream out from an opening in the oven. By the use of slits S_1 and S_2, a sharp linear beam of atoms is obtained. These atoms then pass through a very inhomogeneous magnetic field between the shaped poles of a magnet *M M*. A high degree of non-uniformity in the magnetic field is produced by making one of the pole-pieces of a powerful electromagnet a knife-edge shape and the other flat with a groove cut in it opposite the knife edge [Fig. 6.31 (*b*)]. The lines of force are close together at the knife edge and the field there is much stronger than that at the other pole piece [Fig. 6.31 (*c*)]. The magnetic field is at right angles to the direction of movement of the atoms. Finally the atoms fall on a photographic plate *P*. The whole arrangement is enclosed in an evacuated chamber.

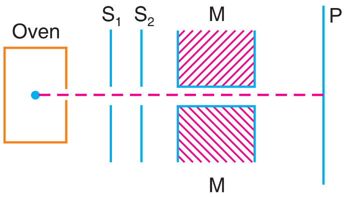


Fig. 6.31 (*a*)

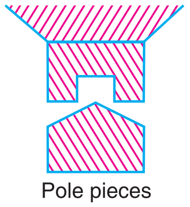


Fig. 6.31 (*b*)

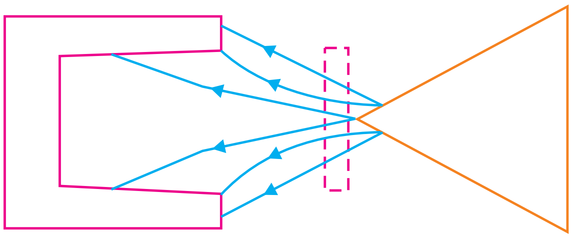


Fig. 6.31 (*c*)

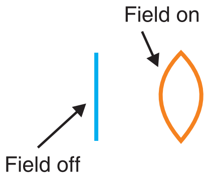
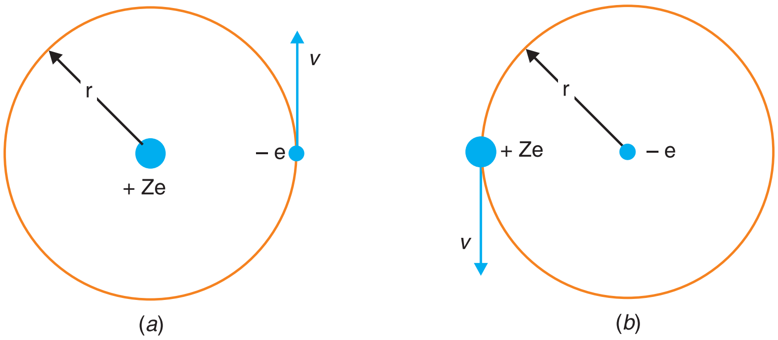


Fig. 6.31 (*d*)

With no field, the beam produces a narrow continuous line on the plate [Fig. 6.31 (*d*)]. In terms of the vector atom model, those atoms, with electron spins directed parallel to the magnetic field, will experience a force in one direction, whereas those with oppositely directed spins will experience a force in the opposite direction. According to this, the beam of atoms should split into two beams in its passage through the inhomogeneous magnetic field. This splitting of the beam into two parts of approximately equal intensity was actually observed in these experiments. On applying the inhomogeneous magnetic field, it was found that the stream of silver atoms splits into two separate lines [Fig. 6.31 (*d*)]. Knowing dB/dx, L, V and α, μ was calculated. It was found that each silver atom had a magnetic moment of one *Bohr magneton* in the direction of the field.

6.21 Spin-Orbit Coupling

The fine structure doubling of spectral lines may be explained on the basis of a magnetic interaction between the spin and orbital angular momenta of atomic electrons. This spin-orbit coupling can be understood in terms of a straight forward classical model. In its orbital motion, the electron goes in a closed path round the nucleus of charge Ze. When viewed from the electron, it is equivalent to the nucleus of charge Ze going round the electron in a closed path [Fig. 6.32]. At the point where the electron is situated, a magnetic induction is produced because of the nucleus going round in a closed path. The interaction between the electron's spin magnetic moment and this magnetic field leads to the phenomenon of spin-orbit coupling. The magnetic potential energy associated with the spin-orbit interaction is



(*a*) (*b*)

Fig. 6.32

$$\Delta E_s = -\vec{\mu}_s \cdot \vec{B} = -\mu_s\ B\cos\theta$$

where θ is the angle between $\vec{\mu}_s$ and $\vec{B}$. The quantity $\mu_s \cos\theta$ is the component of $\vec{\mu}_s$ parallel to *B*. In the case of the spin magnetic moment of the electron this component is $\mu_{sz} = \mu_B$, the Bohr magneton.

$$\mu_s \cos\theta = \mu_{sz} = \mu_B = \pm\frac{e(h/2\pi)}{2m}$$

$$\therefore \qquad \Delta Es = \pm\mu_B B.$$

Depending upon the orientation of its spin vector, the energy of an electron in a given atomic quantum state will be higher or lower by $\mu_B B$ than its energy in the absence of spin-orbit coupling. The result is the splitting of every quantum state into two separate substates. Hence every spectral line splits into two component lines.

Notice that there is one exception. When an electron is in the *s* state, $l = 0$, and as a result the magnetic field at the site of the electron is zero. Since $B = 0$, ΔE_s must also be zero, and there is no fine structure splitting for this state.

EXAMPLE. *Calculate the wavelength separation of the fine structure resulting from the spin-orbit interaction within the hydrogen atom.*

SOL.

$$\left.\begin{array}{l}\text{Magnetic field at the}\\ \text{centre of a circular wire}\\ \text{loop of radius } r \text{ that car-}\\ \text{ries a current } i\end{array}\right\} = B = \frac{\mu_0 i}{2r}.$$

The orbiting electron "sees" itself circled *f* times per second by the proton of charge +*e* that is the nucleus.

$$\left.\begin{array}{l}\therefore\ \text{The magnetic field}\\ \text{resulting from the orbital}\\ \text{motion of the electron}\\ \text{around the nucleus}\end{array}\right\} = \frac{\mu_0\ ef}{2r}$$

For an atom in the ground state, $r_0 = 0.527 \times 10^{-10} m$.

$$f = \frac{v_0}{2\pi\ r_0} = \frac{2.188\times10^6}{2\pi\ (0.527\times10^{-10})} = 6.6\times10^{15}\ s^{-1}.$$

$$\left.\begin{array}{l}\text{Magnetic field exper-}\\ \text{ienced by the electron}\end{array}\right\} = B = \frac{(4\pi \times 10^{-7})\,(1.6 \times 10^{-19})\,(6.6 \times 10^{15})}{2 \times (0.527 \times 10^{-10})} = 12.58\ T.$$

$$\left.\begin{array}{l}\text{The energy of separa-}\\ \text{tion of split lines}\end{array}\right\} = E_s = \frac{hc}{\lambda} - \frac{hc}{\lambda'} = \frac{hc\,\Delta\lambda}{\lambda\,\lambda'} \approx \frac{hc\,\Delta\lambda}{\lambda^2}$$

$$\therefore \qquad \Delta\lambda = \frac{\lambda^2\,E_s}{hc}$$

Now $E_s = 2\,\mu_B\,B = 2\,(9.27 \times 10^{-24})\ 12.58 = 2.332 \times 10^{-22}\,J.$

For example, if $\lambda \approx \lambda' = 4000$ Å, then

$$\Delta\lambda = \frac{(4 \times 10^{-7})^2\ (2.332 \times 10^{-22})}{(6.627 \times 10^{-34})\,(3 \times 10^{8})}\ m = 1.876\ \mathring{A}.$$

6.22 Optical Spectra

The outer (valence) electrons of an atom determine the chemical and *optical* properties of their atoms. The electrons in the closed inner orbits do not take part in the emission of spectral lines.

Spectral terms. Atoms are divided into two main categories, *viz*., one electron system and many-electron system.

(*i*) The alkali metals (such as lithium, sodium and potassium) have a single valence electron outside completely filled subshells. The valence electron in the alkali metals therefore behaves much like the orbiting electron in the hydrogen atom. Hence the alkali metals have a hydrogen-like spectra. Their spectrum is also referred to as the *one-electron spectra.*

(*ii*) In the many-electron system, the atoms have more than one valence or *optical electron* outside completely filled subshells. Hence they become effective in fixing the spectral properties. For example, the alkaline earths belonging to the two-electron system have spectra which are similar among themselves.

Spectral notation. The states of the atom, in which the values of its L vector are 0, 1, 2, 3, 4, 5 are represented by the capital letters, S, P, D, F, G, H, etc., respectively. The value of the total angular momentum of the atom J is written as a subscript at the lower right of the letter representing the particular L value of the atomic state. The multiplicity of the total spin (S) is written as a superscript at the upper left of letter. If S is the total spin, the multiplicity is equal to $2S + 1$.

Examples: (*i*) A state with $L = 1,\ S = \frac{1}{2}$ and $J = 3/2$ would be written ${}^2P_{3/2}$ and read "doublet P three halves". (Since $S = \frac{1}{2}$, the multiplicity of the state $= 2 \times \frac{1}{2} + 1 = 2$).

(*ii*) A state with $L = 2$, $S = 1$ and $J = 2$ would be written ${}^3D_2$ and read "triplet D two".

(*iii*) The spectral terms corresponding to $L = 1$ and $\frac{1}{2}$ are written as ${}^2\boldsymbol{P}_{1/2}$ and ${}^2\boldsymbol{P}_{3/2}$. (Since $L = 1$, the capital letter which represents the state is P; the multiplicity of the state $= 2 \times \frac{1}{2} + 1 = 2$; this is put as a superscript; $J = L \pm S$; $J = 1 + \frac{1}{2} = \frac{3}{2}$ or $1 - \frac{1}{2} = \frac{1}{2}$. $\frac{1}{2}$ or $\frac{3}{2}$ are put as subscripts).

The selection rules. An electron cannot jump from one energy level to all other energy levels. A transition of an electron between two levels is possible only if certain rules called *selection rules* are satisfied. For the vector atom model, three selection rules have been devised.

(*i*) The selection rule for L is $\Delta L = \pm 1$ *i.e.*, only those lines are observed for which the value of L changes by ±1. For example, L can change from 0 to 1 ($\Delta L = 1$) or from 1 to 0 ($\Delta L = -1$). L cannot change from 0 to 2 ($\Delta L = +2$) or from 2 to 0 ($\Delta L = -2$). In symbols, a transition is possible between S and P levels ($\Delta L = \pm 1$), but not possible between S and D levels ($\Delta L = \pm 2$).

(*ii*) The selection rule for J is $\Delta J = \pm 1$ or 0. But $0 \to 0$ is excluded.

(*iii*) The selection rule for S is given by $\Delta S = 0$.

Intensity Rules. Whether an allowed transition is weak or strong is determined by what are known as *intensity rules*. They are:

(*i*) Transitions for which L and J change in the same way (*i.e.*, $\Delta J = \Delta L$) are strong. For other changes in L and J ($\Delta L \neq \Delta J$) we get weak transitions.

(*ii*) Transitions for which L and J increase (*i.e.*, $L \to L + 1$ and $J \to J + 1$) are less intense than those for which L and J decrease (*i.e.*, $L \to L - 1$ and $J \to J - 1$).

(*iii*) Transitions for which changes in L and J are opposite (*i.e.*, $\Delta L = -\Delta J$) are forbidden. Symbolically, these intensity rules are written as

$$\begin{array}{lll} \Delta L = -1, & \Delta J = -1 & \text{strongest} \\ \Delta L = -1, & \Delta J = \ \ 0 & \text{less intense} \\ \Delta L = +1, & \Delta J = +1 & \text{weak} \\ \Delta L = +1, & \Delta J = 0 & \text{very weak} \end{array}$$

$$\left.\begin{array}{ll} \Delta L = -1, & \Delta J = +1 \\ \Delta L = +1, & \Delta J = -1 \end{array}\right\} \text{Forbidden.}$$

The Interval Rule. Lande discovered a rule regarding the interval in frequency between the different levels constituting a multiplet. It states that *the frequency interval between two levels with total angular momenta* ($J + 1$) *and J respectively is proportional to* ($J + 1$).

Fine structure of the sodium D-line

Ten out of the eleven electrons of the normal sodium atom are interlocked in closed shells. They contribute nothing to the angular momentum of the atom. We have to consider only the states of eleventh optical electron in discussing the spectrum of neutral sodium.

The *D*-line belongs to principal series. Lines of the principal series are due to transitions from a P state to the S state. For the upper P state, $L = 1$, $J = L \pm S = 3/2$ or $1/2$. Hence the two possible terms are: $^2P_{3/2}$ and $^2P_{1/2}$. For the lower S state, $L = 0$, $J = \dfrac{1}{2}$, so that only one term $^2S_{1/2}$ is possible. Fig. 6.33 shows the two possible transitions between the two terms of the P state and the single term of the S state. They are : (*i*) $^2P_{1/2} \to {}^2S_{1/2}$ which results in the D_1 line of wavelength 5896 Å and (*ii*) $^2P_{3/2} \to {}^2S_{1/2}$ giving the D_2 line of wavelength 5890 Å. Now applying the selection rules $\Delta L = \pm 1$ and $\Delta J = \pm 1$ or 0 (excluding $0 \to 0$), both the transitions are allowed. This explains the doublet fine structure of the sodium D line.

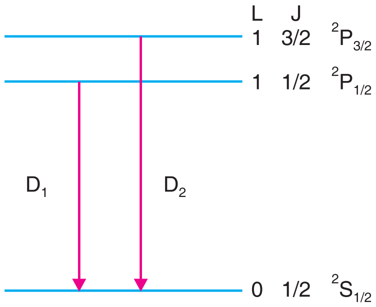


Fig. 6.33

Hyperfine structure. Atomic nuclei have radii of the order of 10^{-4} Å which are small compared with typical distances of an electron from the nucleus (≈ 1 Å). The nuclei are also much heavier than electrons. It is therefore a very good approximation to consider the nuclei to be positive point charges of infinite mass. However, high precision experiments reveal the existence of tiny effects of the electronic energy levels, which cannot be explained if the nuclei are considered to be point charges of infinite mass. These effects were first observed by Michelson in 1891 and Fabry and Perot in 1897. They are called *hyperfine effects*, because they produce shifts of the electronic energy levels which are usually much smaller than those corresponding to the fine structure. The hyperfine components of the spectral radiations are observed using a high resolution interferometer. The wavelength change in the hyperfine structure of single isotopes are very much smaller than the electron spin fine structure. Pauli attributed the hyperfine structure to an angular momentum of the nucleus of the atom. The assumption of orbital and spin angular momentum of the electron could explain completely the fine structure of the spectral lines. The hyperfine structure is caused by the properties of the nucleus. The influence of the nucleus is mainly due to (*i*) the isotope effect and (*ii*) the nuclear spin because of which there will be an intrinsic angular momentum.

6.23 Zeeman Effect

Zeeman Effect.

Zeeman effect is a magneto-optical phenomenon discovered by Zeeman in 1896. He observed that if a source of light producing line spectrum is placed in a magnetic field, the spectral lines are split up into components. When the splitting occurs into two or three lines, it is called *normal Zeeman effect* and can be explained quantitatively by classical theory. The splitting of a spectral line into more than three components in ordinary *weak magnetic fields* is called *anomalous Zeeman effect.* This cannot be explained by classical theory.

Experimental arrangement for the normal Zeeman effect. The arrangement is shown in Fig. 6.34. *MM* is an electromagnet capable of producing a very strong magnetic field. Its conical pole-pieces *PP* have longitudinal holes drilled through them. A source of light (*L*) emitting line spectrum (say, a sodium vapour lamp) is placed between the pole-pieces. The spectral lines are observed with the help of a spectrograph (*S*) of high resolving power. The Zeeman effect may be observed in two ways.

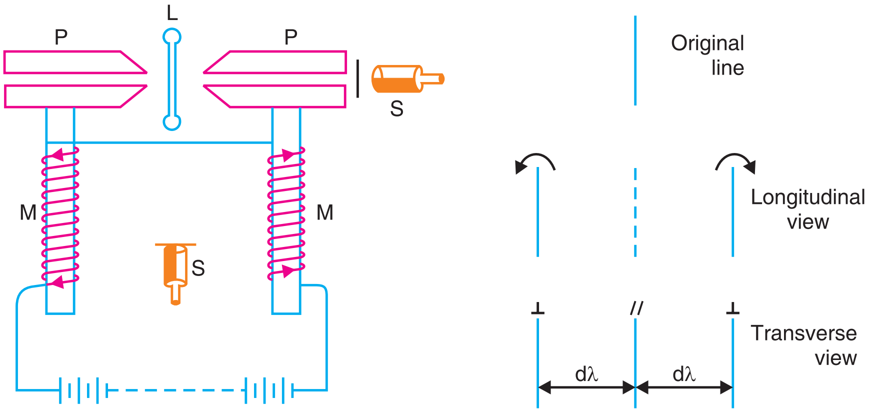


Fig. 6.34

(*i*) The position of the spectral line is noted without applying the magnetic field. The magnetic field is now switched on and the spectral line is viewed *longitudinally* through the hole drilled in the pole-pieces and hence parallel to the direction of the field. It is found that the spectral line is split into two components, one slightly shorter in wavelength and the other slightly longer in wavelength than the original line. The original line is not present. The two components are found to be symmetrically situated about the position of the parent line. Analysing the two lines with a Nicol prism, both the lines are found to be circularly polarised in opposite directions. This is called normal *longitudinal Zeeman effect*.

(*ii*) The spectral line is then viewed *transversely i.e.*, perpendicular to the direction of the magnetic field. In this case, the single spectral line is *split* up into three components when the magnetic field is applied. The central line has the same wavelength as the original line and is plane polarised with vibrations parallel to the field. The outer lines are symmetrically situated on either side of the central line. The displacement of either outer line from the central line is known as the *Zeeman shift*. The two outer lines are also plane polarised having vibrations in a direction perpendicular to the field. This is called *normal transverse Zeeman effect*.

Lorentz classical theory of normal Zeeman effect. The emission of light by a glowing gas is due to the vibratory motion of the electrons. The electrons in the atom execute simple harmonic vibration about the centre of the atom. The frequency of the spectral line is given by the frequency of vibration of the electron. Any linear motion of an electron can be resolved into three components—a linear motion along the magnetic field and two opposite circular motions perpendicular to the field. It is known that no mechanical force acts upon a charge when it moves in the direction of the field. So the linear motion along the field is not affected by the field and hence the frequency of this component is unchanged by the magnetic field.

The other two circular components are affected by the magnetic field, one being retarded and the other being accelerated. Hence the frequency of one of these circular components is increased and that of the other is decreased by the same amount.

In the longitudinal view, the unaltered linear component along the field gives no light waves along the direction of observation due to the transverse nature of light vibrations. Hence the original line is not observed. The two altered circular components at right angles to the field produce circularly polarised light waves. This explains the longitudinal Zeeman pattern.

In the transverse view, the unaffected linear component sends out vibrations parallel to the field and hence perpendicular to the direction of observation. This gives the central line of the triplet occupying the same position as the original line and plane polarised with vibrations parallel to the field. The two circular vibrations which are altered in frequency send out vibrations perpendicular to the field and when viewed sideways appear as plane polarised vibrations in opposite senses.

Expression for the Zeeman shift. Consider an electron in the atom moving in a circular orbit of radius r with a linear velocity v and angular velocity ω. Let e be the charge of the electron and m its mass.

$$\left.\begin{array}{l}\text{The centripetal force on}\\ \text{the electron towards the}\\ \text{centre in the absence of}\\ \text{the magnetic field}\end{array}\right\} = F = \frac{mv^2}{r} = m\omega^2 r \qquad \text{...(1)}$$

Now let an external magnetic field of flux density B be applied in a direction perpendicular to the plane of the orbits of the two circular components and directed towards the reader. Then an additional radial force of magnitude Bev acts on the electron. The direction of this force will be outwards from the centre for clockwise motion but inwards towards the centre for anticlockwise motion (Fig. 6.35). The resulting complex motion of the electron subjected to an additional radial force is called *Larmor precession.* This produces a change in the angular velocity without any change in the form of the orbit.

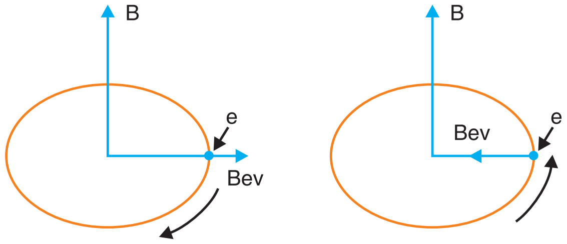


Fig. 6.35

Let $\delta\omega$ be the change in angular velocity caused by the field. For the circular motion in the *clockwise* direction, the additional radial force is directed away from the centre.

$$\therefore \qquad F - Bev = m\,(\omega + \delta\omega)^2\, r \qquad \text{...(2)}$$

$$m\omega^2 r - m\,(\omega + \delta\omega)^2\, r = Be\omega r$$

$$\text{or} \qquad -2mr\omega\,\delta\omega = Be\omega r \qquad \text{neglecting } (\delta\omega)^2$$

$$\text{or} \qquad \delta\omega = -Be/2m \qquad \text{... (3)}$$

For the circular motion in the anticlockwise direction, the additional radial force is directed towards the centre.

$$\therefore \qquad F + Bev = m\,(\omega + \delta\omega)^2\, r$$

$$\text{or} \qquad \delta\omega = +Be/2m \qquad \text{...(4)}$$

The two cases can be combined into the equation

$$\delta\omega = \pm\frac{Be}{2m} \qquad \text{...(5)}$$

If ν = frequency of vibration of the electron,

$$\omega = 2\pi\nu;\ \delta\omega = 2\pi\ \delta\nu \quad \text{or} \quad \delta\nu = \frac{\delta\omega}{2\pi}.$$

$$\left.\begin{array}{l}\text{Change in frequency}\\ \text{of the spectral line}\end{array}\right\} = \delta\nu = \pm\frac{Be}{4\pi m} \qquad \text{...(6)}$$

If ν and λ are the frequency and wavelength of the original line,

$$\nu = \frac{c}{\lambda} \quad \text{or} \quad \delta\nu = \frac{-c}{\lambda^2}\,\delta\lambda.$$

$$\therefore \qquad \text{The Zeeman shift} = \delta\lambda = \pm\frac{Be\lambda^2}{4\pi mc} \qquad \text{...(7)}$$

Note

Taking a spectral line of known wavelength λ, and applying a magnetic field B, the Zeeman shift $d\,\lambda$ is measured.

Then $\frac{e}{m} = \left(\frac{4\pi c}{B\lambda^2}\right) d\lambda$ can be calculated. e/m is found to be $1.757 \times 10^{11}\ C\ \text{kg}^{-1}$ which is in agreement with the value of e/m of the electron obtained from Thomson's experiment.

Example 1. *Calculate the wavelength separation between the unmodified line of wavelength 6000 Å and the modified lines when a magnetic induction of 1 Wbm^{-2} is applied, in normal Zeeman effect.*

Sol. We have

$$d\lambda = \frac{Be\lambda^2}{4\pi mc} = \frac{1 \times (1.6 \times 10^{-19})\,(6000 \times 10^{-10})^2}{4\pi\,(9.1 \times 10^{-31})\,(3 \times 10^8)}$$

$$= 0.168 \times 10^{-10}\ m = 0.168\ \text{Å}.$$

Example 2. *Calculate the wavelength separation between the two component lines which are observed in the normal Zeeman effect. The magnetic field used is 0.4 weber/m^2; the specific charge = 1.76×10^{11} C kg^{-1} and λ = 6000 Å.*

Sol. The wavelength separation

$$\lambda_1 - \lambda_2 = 2\,d\lambda = \frac{Be\lambda^2}{2\pi mc}$$

Here, $B = 0.4 \text{ weber/m}^2;\ e/m = 1.76 \times 10^{11}\ C\,\text{kg}^{-1}$

$\lambda = 6000 \times 10^{-10}\text{m};\ c = 3 \times 10^8 \text{ ms}^{-1};$

$$\therefore \quad \lambda_1 - \lambda_2 = \frac{0.4 \times (1.76 \times 10^{11})\,(6000 \times 10^{-10})^2}{2\pi\,(3 \times 10^8)}\ m$$

$$= 1.335 \times 10^{-11} \text{ m} = 0.1335 \text{ Å}$$

Example 3. *The red line of cadmium splits into three components separated by* 120 *MHz when the source is placed in a magnetic field of flux density* 8.6 *mT, the light being examined in a direction perpendicular to the magnetic field. Calculate the ratio of charge to mass* (*e*/*m*) *of the electron.*

Sol. We have, $\delta\nu = Be/4\,\pi m.$

$$\therefore \quad \frac{e}{m} = \frac{4\pi\,\delta\nu}{B} = \frac{4\pi \times (1.2 \times 10^8)}{8.6 \times 10^{-3}} = 1.76 \times 10^{11}\ C\ \text{kg}^{-1}$$

6.24 Larmor's Theorem

Statement. *The effect of a magnetic field on an electron moving in an orbit is to superimpose on the orbital motion a precessional motion of the entire orbit about the direction of the magnetic field with angular velocity* ω *given by* $\omega = Be/2m$.

Explanation. Fig. 6.36 shows two positions of the vector $\vec{l}$ as it precesses about the magnetic field at constant inclination and the corresponding positions of the electronic orbit. Just as a mechanical top precesses in a gravitational field, an electron in an orbit precesses in a magnetic field. This is called Larmor precession.

Change in K.E. due to Larmor precession. In the magnetic field, the angular velocity of the atomic system changes by an amount $Be/2m$. Let ω_0 be the original angular velocity. Then change in K.E.

$$= \frac{1}{2} mr^2 \left(\omega_0 + \frac{Be}{2m}\right)^2 - \frac{1}{2} mr^2\ \omega_0{}^2$$

$$= \frac{1}{2} mr^2\ \omega_0^2 \left(1 + \frac{Be}{2m\omega_0}\right)^2 - \frac{1}{2} mr^2\ \omega_0{}^2$$

$$= \frac{1}{2} mr^2\ \omega_0^2 \left(1 + \frac{Be}{m\omega_0}\right) - \frac{1}{2} mr^2\ \omega_0^2$$

$$= mr^2\ \omega_0^2\ \frac{Be}{2m\omega_0} = \text{L}\ \frac{Be}{2m}$$

where **L** = original angular momentum = $mr^2\ \omega_0$.

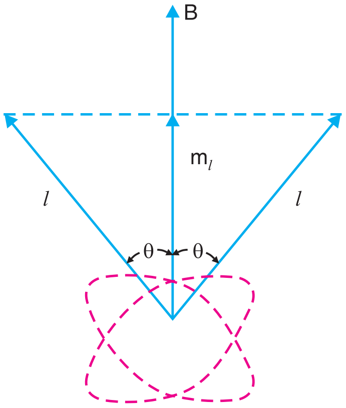


Fig. 6.36

6.25 Quantum Mechanical Explanation of the Normal Zeeman Effect

Debye explained the normal Zeeman effect without taking into account the concept of electron spin. If we neglect the spin motion of the electron, then,

$$\left.\begin{array}{l}\text{the orbital angular}\\ \text{momentum of the}\\ \text{electron}\end{array}\right\} = \mathbf{L} = \frac{lh}{2\pi} \qquad \text{...(1)}$$

$$\text{and magnetic moment } = \mu_l = \frac{lh}{2\pi}\ \frac{e}{2m} = \frac{e}{2m}\ \mathbf{L} \qquad \text{...(2)}$$

In the presence of an external magnetic field of flux density **B**, the vector **1** precesses around the direction of the magnetic field as axis. The precession is known as Larmor precession.

The frequency of Larmor precession = ω = $Be/2m$ ...(3)

Fig. 4.36 shows two positions of the vector **1**, as it precesses about the magnetic field at constant inclination and the corresponding positions of the electronic orbit. The additional energy of the electron due to this precessional motion

$$\Delta E = \mu_l B \cos\theta = \left(\frac{e}{2m}\frac{lh}{2\pi}\right) B\cos\theta \qquad \left(\because\ \mu_l = \frac{e}{2m}\frac{lh}{2\pi}\right).$$

$$= \frac{Be}{2m}\frac{h}{2\pi} l \cos\theta$$

But $Be/2m = \omega$ and $l \cos\theta$ = projection of **1** on **B** = m_l

$$\therefore \qquad \Delta E = m_l \frac{eh}{4\pi m} B = m_l \omega \frac{h}{2\pi} \qquad ...(4)$$

Now, m_l can have $(2l + 1)$ values from $+ l$ to $- l$. Therefore, an external magnetic field will split a single energy level into $(2l + 1)$ levels. The *d*-state ($l = 2$) is split into 5 sub-levels and the *p*-state ($l = 1$) is split into 3 sublevels (Fig. 6.37).

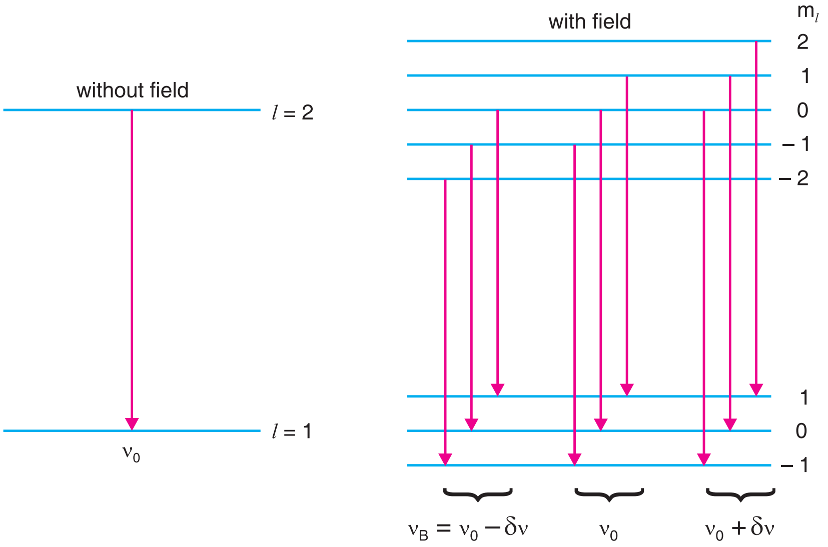


Fig. 6.37

Let E_o' represent the energy of the level $l = 1$ in the absence of the magnetic field and E_B' represent the energy of this level in the presence of magnetic field. Then,

$$E_B' = E_o' + \Delta E' = E_o' + m_l' \frac{eh}{4\pi m} B \qquad ...(5)$$

Similarly, if E_o'' and E_B'' represent the energies of the level $l = 2$ without and with the magnetic field respectively, then,

$$E''_B = E''_o + \Delta E'' = E''_o + m''_l \frac{eh}{4\pi m} B \qquad ...(6)$$

The quantity of energy radiated in the presence of magnetic field is

$$E''_B - E'_B = (E''_o - E'_o) + (m''_l - m'_l)\frac{eh}{4\pi m} B$$

$$h\nu = h\nu_o + \Delta m_l \frac{eh}{4\pi m} B$$

or

$$\nu = \nu_o + \Delta m_l \frac{eB}{4\pi m} \qquad ...(7)$$

where ν = frequency of the radiation emitted with the magnetic field and ν_o = frequency of the radiation in the absence of the magnetic field. The selection rule for m_l is $\Delta m_l = 0$ or ± 1.

Hence we have three possible lines,

$$\nu_1 = \nu_o \text{ for } \Delta m_l = 0 \qquad ...(8)$$

$$\nu_2 = \nu_0 + \frac{eB}{4\pi m} \text{ for } \Delta m_l = +1 \qquad ...(9)$$

and

$$\nu_3 = \nu_o - \frac{eB}{4\pi m} \text{ for } \Delta m_l = -1 \qquad ...(10)$$

Fig. 6.37 represents the normal Zeeman effect. Although there are nine possible transitions, they are grouped into only three different frequency components as indicated by equations (8), (9) and (10). For three transitions in a bracket, change in the value of Δm_l is the same and hence they represent same change of energy and a single line.

6.26 Anomalous Zeeman Effect

This can be explained only by using the idea of the spin of the electron. With the introduction of spin, we have two angular momentum vectors **1** and **s** associated with each electron. The total angular momentum vector $\mathbf{j} = \mathbf{1} + \mathbf{s}$ ...(1)

The magnetic moment due to orbital motion

$$= \mu_l = \mathbf{l} \frac{eh}{4\pi m} \qquad ...(2)$$

μ_l is directed oppositely to **l** because of the negative charge of the electron.

Similarly, the magnetic moment due to the spin of the electron

$$\mu_s = 2\text{s} \frac{eh}{4\pi m} \qquad ...(3)$$

μ_s is oppositely directed to **s** because of the negative charge of the electron. The relationships between the magnetic moments and the angular momenta are shown in Fig. 6.38.

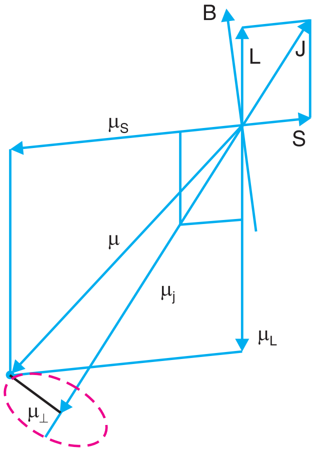


Fig. 6.38

The resultant magnetic moment μ is not along **j**. Since **1** and **s** precess about **j**, μ_l and μ_s must also precess about **j**.

To find out the resultant magnetic moment of the electron, each of these vectors μ_l and μ_s is resolved into two components, one along **j** and the other perpendicular to it. The value of the perpendicular component of each vector, averaged over a period of the motion will be zero, since it is constantly changing direction. The effective magnetic moment of the electron will be

μ_j = component of μ_l along the direction of **j** + component of μ_s along the direction of **j**

$$= \frac{eh}{4\pi m} \mathbf{1} \cos(\mathbf{1}, \mathbf{j}) + \frac{eh}{4\pi m} 2\mathbf{s} \cos(\mathbf{s}, \mathbf{j})$$

$$= \frac{eh}{4\pi m}\left[\mathbf{1}\cos\,(\mathbf{1},\mathbf{j}) + \mathbf{2s}\cos\,(\mathbf{s},\mathbf{j})\right]$$

But according to cosine law,

$$\cos\,(\mathbf{1},\mathbf{j}) = \frac{l^2 + j^2 - s^2}{2lj}$$

and

$$\cos\,(\mathbf{s},\mathbf{j}) = \frac{s^2 + j^2 - l^2}{2sj}$$

Hence,

$$\mu_j = \frac{eh}{4\pi m}\left[\frac{l^2 + j^2 - s^2}{2j} + \frac{s^2 + j^2 - l^2}{j}\right]$$

$$= \frac{eh}{4\pi m}\left[\frac{3j^2 + s^2 - l^2}{2j}\right]$$

$$= \frac{eh}{4\pi m}\, j\left[1 + \frac{j^2 + s^2 - l^2}{2j^2}\right]$$

Writing $j^2 = j\,(j+1)$ and so on,

$$\mu_j = \frac{eh}{4\pi m}\, j\left[1 + \frac{j(j+1) + s(s+1) - l(l+1)}{2j(j+1)}\right]$$

The quantity $1 + \dfrac{j(j+1) + s(s+1) - l(l+1)}{2j(j+1)} = g$

is called the Lande g factor.

Hence

$$\mu_j = \frac{eh}{4\pi m}\, gj.$$

If the atom is placed in a weak magnetic field, the total angular momentum vector **j** precesses about the direction of the magnetic field as axis.

The additional energy ΔE due to the action of the magnetic field on this atomic magnet is

$$\Delta E = \mu_j\, B\quad \cos\,(j, B) = \frac{eh}{4\pi m}\, g\; jB\quad \cos\,(j, B).$$

But $j\cos\,(j, B)$ = the projection of the vector **j** on the direction of the magnetic field = m_j.

Hence,

$$\Delta E = \frac{eh}{4\pi m}\, Bgm_j.$$

The quantity $\dfrac{eh}{4\pi m}B$ is called a *Lorentz unit*. It is a unit of energy used for expressing the splitting of the energy levels in a magnetic field.

Since m_j has $(2j + 1)$ values, a given energy level is split up into $(2j + 1)$ sublevels with the application of magnetic field. When this m_j is subjected to the selection rule $\Delta m_j = 0$ or ± 1, we get the transitions shown in Fig. 6.39 for the sodium D-lines.

The ground state $^2S_{1/2}$ splits into two sublevels. Here,

$$l = 0,\ s = \frac{1}{2};\quad j = \frac{1}{2};\quad \text{hence,}$$

$$g = 1 + \frac{\frac{1}{2}\times\frac{3}{2} + \frac{1}{2}\times\frac{3}{2}}{2\times\frac{1}{2}\times\frac{3}{2}} = 2.$$

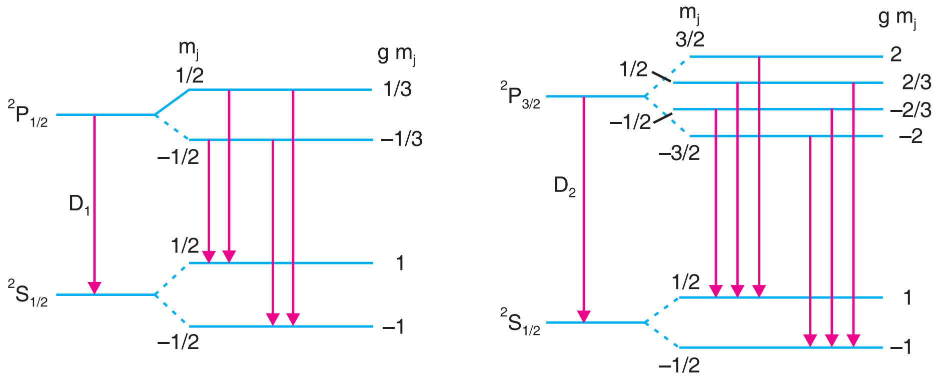


Fig. 6.39

Since m_j can have the values $\frac{1}{2}$ and $-\frac{1}{2}$, gm_j can have the values +1 and –1. The following table (Table 6.3) gives the values for the quantum numbers necessary for the determination of the splitting factor gm_j for each of the energy levels of the sodium D-lines.

The longer wavelength component $^2P_{1/2} \rightarrow {}^2S_{1/2}$ splits into four lines. The shorter wavelength component $^2P_{3/2} \rightarrow {}^2S_{1/2}$ splits into six lines.

Thus the introduction of electron spin has led to complete agreement between the experimental results and the theory of the anomalous Zeeman effect.

TABLE 6.3

State	*l*	*s*	*j*	*g*	m_j	$m_j g$
$^2S_{1/2}$	0	$\frac{1}{2}$	$\frac{1}{2}$	2	$\frac{1}{2}, -\frac{1}{2}$	1, –1
$^2P_{1/2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{2}{3}$	$\frac{1}{2}, -\frac{1}{2}$	$\frac{1}{3}, -\frac{1}{3}$
$^2P_{3/2}$	1	$\frac{1}{2}$	$\frac{3}{2}$	$\frac{4}{3}$	$\frac{3}{2}, \frac{1}{2}$ $-\frac{1}{2}, \frac{-3}{2}$	$2, \frac{2}{3}$ $-\frac{2}{3}, -2$

6.27 Paschen-Back Effect

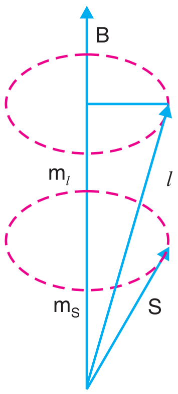


Fig. 6.40

Paschen and Back found that whatever be the anomalous Zeeman pattern of a given line in a weak magnetic field, the pattern always approximates the normal Zeeman triplet as the field strength is progressively increased. This reduction may occur either through the coalescence of lines or through the disappearance of certain lines. This transition phenomenon is called Paschen-Back effect.

Explanation. In a strong magnetic field, the coupling between **l** and **s** breaks down, and **j** loses its significance. **l** and **s** are quantised separately. **l** and **s** precess separately about the external magnetic field B independent of each other (Fig. 6.40). The energy change due to the presence of the field will then be made up of two parts, one arising from the precession of **l** about **B** and the other from the precession of **s** about **B.**

Hence

$$\Delta E = (\Delta E)_l + (\Delta E)_s$$

$$= B\,\frac{eh}{4\pi m}\,[l\cos(l, B) + 2s\cos(s, B)]$$

$$= \frac{eh}{4\pi m} B\,(m_l + 2m_s).$$

In terms of frequency change,

$$\Delta\nu = \frac{eB}{4\pi m} \Delta\,(m_l + 2m_s).$$

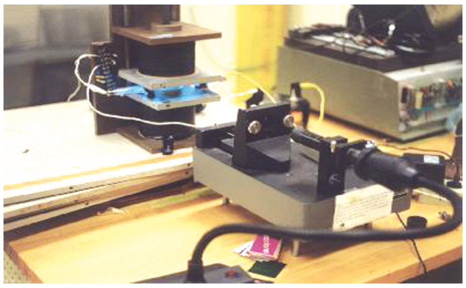
The quantity $(m_l + 2m_s)$ is known as the *strong field quantum number* and is evidently an integer. Now since $\Delta m_l = 0$ or ± 1, $\Delta m_s = 0$, $\Delta(m_l + 2m_s) = 0$ or ± 1.

Hence in a strong magnetic field, a given spectral line will split into three components only and this is the usual characteristic of the normal Zeeman effect.

6.28 Stark Effect

The Stark effect is the electrical analogue of the Zeeman effect. The Stark effect is the splitting of spectral lines due to the action of an external electric field on the radiating substance. Even very strong external electric fields are weak compared to the interatomic fields. Hence the action of electric field on the motion of the atomic electrons can be regarded as small perturbations. Consequently, the Stark line splitting is very minute and can be observed only with instruments having a high resolving power. The lines are split into a series of components (satellites) located, in case of hydrogen, symmetrically on both sides of the original line.

Experimental study. Here the hydrogen atoms emitting spectral lines are subjected to a powerful electric field. The arrangement used by Stark is shown in Fig. 6.41. The canal rays are produced in an ordinary glass discharge tube provided with a perforated cathode *C*. When the pressure in the tube is not very low, discharge takes place between the anode *A* and cathode *C* maintained at a suitable *P.D.* The canal rays stream through the perforations in the cathode and form behind the cathode narrow cylindrical bundles of luminous rays. An auxiliary electrode *F* is placed parallel and close to *C* at a distance of a few millimetres. A very strong electric field of several thousand volts per metre is maintained between *F* and *C*. The effect produced can be studied both transversely [Fig. 6.41 (*a*)] and longitudinally [Fig. 6.41 (*b*)]. Stark observed that the lines in the spectrum emitted by the canal rays of hydrogen were split up into numerous sharp components under the action of the electric field.



Experimental Setup of Stark Effect.

Results. The results obtained with the lines of the Balmer series of the hydrogen spectrum are given below:

(*i*) Every line is split up into a number of sharp components. All hydrogen lines form symmetrical patterns. The pattern depends markedly on the quantum number *n* of the term involved. The number of lines and the total width of the pattern increases with *n*. Thus, the number of components of H_β line is greater than that of the H_α line; similarly, the number of components of H_γ, is greater than that of H_β.

(*ii*) Observation perpendicular to the direction of the electric field (transverse view) shows that the components are polarised, some parallel to the direction of the field and others perpendicular to it.

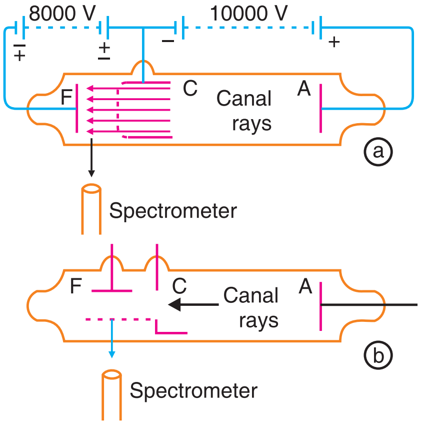


Fig. 6.41

(*iii*) Upto fields of about 10^7 V/m, the resolution increases in proportion to the field strength (E). In this region, we have *linear* or *first order* Stark effect. When E exceeds 10^7 V/m, there are shifts in the line patterns which are proportional to E^2 and we speak of the *second-order* Stark effect.

EXERCISE

1. Describe Rutherford's model of the atom and the evidence that led to it. What are its drawbacks ?

2. Explain the phenomenon of large angle scattering of α-particles and derive the Rutherford's formula for it. Describe how the size of the nucleus and nuclear model have been obtained from the scattering phenomenon.

3. State the postulates of Bohr regarding his atom model. Obtain expressions for the radius and electron-energy of the n^{th} orbit. Explain how Bohr's atom model successfully accounts for the hydrogen spectrum.

4. Give an account of Bohr's theory of hydrogen spectrum taking into account the finite mass of H_2 nucleus.

5. Define the terms : (1) Critical potential, (2) Excitation potential and (3) Ionisation potential. Describe Frank-Hertz experiment for determining the critical potentials. Show how the results of the experiment confirm Bohr's postulates.

6. Distinguish between excitation and ionisation potential of a gas. Describe with necessary theory Davis and Goucher's method for determining the critical potentials of a gas.

7. Give an account of Bohr-Sommerfeld model of elliptical electron orbits of hydrogen atom. How does it account for the fine structure of hydrogen ?

8. Discuss the general characteristics of Sommerfeld's elliptical orbits and show that the total energy of an electron moving in Sommerfeld's orbits of the same total quantum number is the same and is identical with that of corresponding Bohr orbits. How is the result modified by introducing correction for relativistic variation of electron mass ?

9. Describe the vector model of the atom and explain the different quantum numbers associated with it. Write down the electron configuration for *Cu* (29) employing modern symbolism and explain it.

10. State and explain Pauli's exclusion principle as applied to electrons in atoms. Describe how this principle assists in the interpretation of the periodic system of the elements.

11. What is Pauli's exclusion principle ? On the basis of this principle explain the configuration of electrons in atoms.

12. Describe the Stern and Gerlach experiment and indicate the importance of the results obtained.

13. How will you explain the D_1 and D_2 doublet of sodium spectrum on the basis of electron spin?

14. What is Zeeman effect ? Describe the experimental arrangement for studying the Zeeman effect. Use classical ideas to explain normal Zeeman effect. Show that the Zeeman shift $d\lambda = \pm \frac{Be\lambda^2}{4\pi mc}$.

15. Outline the quantum theory of Zeeman effect. Illustrate your answer with specific reference to Sodium *D*-line.

16. Explain clearly the phenomenon of anomalous Zeeman effect and Paschen-Back effect. Describe the spectral patterns expected for the yellow lines of sodium according to these two effects.

17. Derive an expression for Lande's splitting factor and explain the anomalous Zeeman effect of sodium doublet lines D_1 and D_2 with its help.

18. The well known *D*-lines of sodium come from the transition $3P\ {}^2P_{1/2,\ 3/2} \rightarrow 3S\ {}^2S_{1/2}$. Work out the Zeeman patterns for the levels, including the *g* values, and from this deduce the Zeeman patterns of the resulting spectral lines. Distinguish between Paschen-back, normal and anomalous Zeeman effect.

19. With gold foil (Z = 79), 180° scattering is observed with α-particles of velocity up to $1.6 \times 10^7\ ms^{-1}$. Deduce the upper limit for radius of the nucleus. [**Ans.** $4.26 \times 10^{-14} m$]

20. Explain why electrons are not effective in scattering α-particles and protons.

[**Ans.** α-particle is ≈ 7000 times heavier than an electron. The electrons, because of their little mass, are unable to deviate the α-particles and protons appreciably.]

21. Calculate the closest distance of approach of an α-particle of energy 5 *MeV* shot at a gold nucleus. [**Ans.** $4.54 \times 10^{-14} m$]

22. The Rutherford scattering formula fails to agree with the data at very small scattering angles. Can you think of a reason ?

[**Ans.** Small θ implies a large impact parameter. In this case, the full nuclear charge of the target atom is partially screened by its electrons.]

23. A 5–*MeV* alpha particle approaches a gold nucleus with an impact parameter of $2.6 \times 10^{-13} m$. Through what angle will it be scattered ? [**Ans.** 10°]

24. For equal numbers of incident particles, determine the ratio of protons to alpha particles of the same energy that will be scattered by angles greater than 90° from a gold foil. [**Ans.** 0.25]

25. In the Bohr theory of the hydrogen atom, the electron is in constant motion. How is it possible for such an electron to have a negative amount of energy ?

[**Ans.** The P.E. is more negative than the K.E. is positive]

26. The energy of the electron in the n^{th} orbit in hydrogen atom is negative. Explain.

[**Ans.** An electron revolving round the nucleus in a hydrogen atom has a negative P.E. by virtue of its attraction towards the nucleus. The electron has K.E. (which is positive) by virtue of its motion. The P.E. is greater in magnitude than the K.E., so that the net energy is negative. The '*negative*' energy signifies that the electron cannot escape from the atom. A positive energy for a nucleon-electron combination would mean that the electron is not bound to the nucleus. Such a combination cannot constitute an atom.]

27. Find the wavelength of the photon emitted when the hydrogen atom goes from $n = 10$ state to the ground state. [**Ans.** 921 Å]

28. Why could not Bohr allow the quantum number n to take on the value $n = 0$?

[**Ans.** If $n = 0$, then $r = 0$, *i.e.*, the electron would have to go through the nucleus which is impossible.]

29. The Rydberg constant for hydrogen is 10967700 m^{-1}. Calculate the short and long wavelength limits of Lyman series. [**Ans.** 911.6 Å, 1215 Å]

30. Calculate the radii of the first, second and third permitted electron Bohr-orbits in a hydrogen atom.

[**Hint**. Radius of the n^{th} Bohr orbit of hydrogen = $\dfrac{n^2h^2\varepsilon_0}{\pi e^2 m} = 0.527\ n^2$ Å.

$r_1 = 0.527$ Å; $r_2 = 2.108$ Å; and $r_3 = 4.743$ Å.]

31. Calculate the value of Rydberg constant, assuming that wavelength of H_α line is 6563 Å.

[**Hint.** $\bar{\nu} = \dfrac{1}{\lambda H_\alpha} = R\left(\dfrac{1}{4} - \dfrac{1}{9}\right) = \dfrac{5R}{36}$

$$\therefore \quad \frac{1}{6563 \times 10^{-10}} = \frac{5R}{36} \quad \text{or} \quad R = \frac{36}{5 \times 6563 \times 10^{-10}} = 1.097 \times 10^7\ m^{-1}]$$

32. If the Rydberg constant is 1.097×10^7 m^{-1}, what are the wavelengths of the first three lines of the Paschen series ?

[**Hint.** The formula for Paschen series is,

$$\bar{\nu} = R\left(\frac{1}{3^2} - \frac{1}{n^2}\right);\ n = 4, 5, 6, \ldots..$$

$$\bar{\nu}_1 = \frac{1}{\lambda_1} = 1.097 \times 10^7 \left(\frac{1}{3^2} - \frac{1}{4^2}\right) = 1.097 \times 10^7 \times \frac{7}{144};$$

$$\therefore \quad \lambda_1 = \frac{144}{1.097 \times 10^7 \times 7}\ m = 18750\ Å;\ \lambda_2 = 12810\ Å;\ \lambda_3 = 10930\ Å]$$

33. When hydrogen was bombarded in Frank-Hertz experiment by 10.21 *eV* and 12.10 *eV* electrons, emission of three spectral lines was observed. Calculate their wavelengths. [**Ans.** 1026 Å, 1216 Å, 5672 Å]

34. If excitation energy is provided by inelastic collisions amongst the molecules, calculate the temperature at which mercury vapour will emit ultraviolet light of 2586 Å. [**Ans.** 5.5×10^4 *K*]

35. Show that the second ionisation potential of helium is 54.34 volts, given $e = 1.6 \times 10^{-19}$ *C*; $m = 9.11 \times 10^{-31}$ kg; $h = 6.62 \times 10^{-34}$ *Js*.

36. A photon of wavelength 1400 Å, is absorbed by cold mercury vapour and two other photons are emitted. If one of them is of 1850 Å line, what is the wavelength of the other photon ?

[**Hint.** *Energy absorbed* $E = \frac{hc}{\lambda} = \frac{(6.62 \times 10^{-34})\,(3 \times 10^{8})}{1400 \times 10^{-10}}$

$$= \frac{(6.62 \times 10^{-34})\,(3 \times 10^{8})}{(1400 \times 10^{-10}) \times (1.6 \times 10^{-19})}\, eV = 8.867\ eV$$

1850 Å absorbs 6.71 eV; balance 2.157 eV emits a line of wavelength 5755 Å]

37. Explain Bohr's theory of the spectra of hydrogen and of ionised helium atoms and determine from the expressions of the Rydberg's constant the ratio of the mass of electron to that of proton. $R_H = 10967800$ m^{-1} and $R_{He} = 10972200$ m^{-1}. [**Ans.** 1/1869]

38. The wavelength of the second line of the Balmer series in the hydrogen spectrum is 4861 Å. Calculate the wavelength of the first line. [**Ans.** 6563 Å]

39. At what speed must the electron revolve round the nucleus of a hydrogen atom in order that it may not be pulled into the nucleus by electrostatic attraction ? Take the radius of the orbit of the electron as $0.5 \times 10^{10-}$ metre, $m = 9.1 \times 10^{-31}$ kg and $e = 1.6 \times 10^{-19}$ *coulomb*. [**Ans.** $2.25 \times 10^{6}\ ms^{-1}$]

40. The energy levels of a hypothetical one electron atom (not hydrogen) are given below:

n	1	2	3	4	5	∞
$E_n(eV)$	–15.6	–5.3	–3.08	–1.45	–0.80	0

Draw the energy level diagram. Calculate the frequency of photon emitted when the atomic system goes from energy state $n = 2$ to the ground level. [**Ans.** $\nu = 2.5 \times 10^{15}\ Hz$]

41. The ionisation potential of an atom is 14.2 eV. Calculate the series limit in its absorption spectrum.

[**Hint.** $\lambda = \frac{hc}{eV} = \frac{(6.6 \times 10^{-34})\,(3 \times 10^{8})}{(1.6 \times 10^{-19})\,(14.2)}\, m = 871.5\ Å$].

42. In Franck-Hertz experiment, the excitation and ionisation potentials for sodium are 2.1, 3.7 and 5.13 volts respectively. Calculate the velocities of the impinging electrons to excite and ionise the sodium atoms.

[**Hint.** $\frac{1}{2} mv^2 = eV$ or $v = \sqrt{\frac{2eV}{m}}$.

$v_1 = 0.8592 \times 10^{6}\ \text{ms}^{-1}$, $v_2 = 1.140 \times 10^{6}$ ms^{-1} and $v_3 = 1.343 \times 10^{6}$ ms^{-1}].

43. The resonance potentials for mercury are 4.9 and 6.7 V. Calculate the wavelengths of corresponding lines emitted. [**Ans.** 2533 Å; 1852 Å.]

44. Calculate the critical potential in volts corresponding to the excitation of mercury line $\lambda = 5416$ Å. [**Ans.** 2.29 volts.]

45. A critical potential for hydrogen is 13.05 V. Calculate the wavelength of the radiation that will be emitted by a hydrogen atom when bombarded by an electron of corresponding energy.

$$\left[\textbf{Hint.}\ \ \lambda = \frac{hc}{eV} = \frac{(6.62 \times 10^{-34})\,(3 \times 10^{8})}{(1.6 \times 10^{-19})\,(13.05)}\, m = 951.5\ Å.\right]$$

46. Ultraviolet light of 800 Å is incident on hydrogen in a quartz tube. Calculate the kinetic energies with which electrons will be ejected from the hydrogen atoms. Express the results in electron volts. [**Ans.** 1.85 eV; 12.06 eV; 13.94 eV; 14.60 eV.]

47. Calculate the wavelength difference between the second Balmer line ($\lambda = 4800$ Å) of hydrogen and that of heavy hydrogen. Rydberg constant for hydrogen = 10967760 m^{-1} and $R_\infty = 10973740$ m^{-1}. [**Ans.** 1.371 Å]

48. A beam of silver atoms in a Stern-Gerlach experiment, obtained from an oven heated to a temperature of 1500 K, passes through an inhomogeneous magnetic field having a field gradient of 2 weber/m^2/cm perpendicular to the beam. The pole faces are 10 cm long. What is the separation between the two components of the beam on a photographic plate placed at a distance of 50 cm ? [**Ans.** 0.015 cm]

49. Calculate the value of Bohr magneton. $h = 6.6 \times 10^{-34}$ joule second; $m = 9.1 \times 10^{-31}$kg; $e = 1.6 \times 10^{-19}$ coulomb. [**Ans.** 9.253×10^{-24} *SI units*]

50. The experimental value of Bohr magneton is 9.21×10^{-24} *SI units* and Planck's constant $h = 6.6 \times 10^{-34}$ joule-second. Calculate the value of e/m of an electron. [**Ans.** $1.755 \times 10^{11} C\ \text{kg}^{-1}$]

51. Find the precessional frequency of an electron orbit when placed in a magnetic field of $5T$.

[**Ans.** 7×10^{10} *Hz*]

52. Write drawbacks of Bohr's theory [(Barkatullah University, Bhopal, 2011)].

[**Ans.** (1) In Bohr theory, we have two rival theories, *viz*., classical and quantum. The equilibrium is governed by the classical laws, while the emission of radiation is explained by quantum rules.

(2) The theory does not give any information about the relative intensities of different lines.

(3) Bohr's theory could not explain the *fine structure of the spectral lines*.

(4) With the help of Bohr theory, it is difficult to treat dynamical problems of atoms containing more than one valence electron.

(5) Bohr model could not explain how the orbital electrons in an atom are distributed around the nucleus.

(6) Bohr's theory could not explain Zeeman Effect and Stark Effect in which the spectral lines could be split up under the influence of magnetic and electric fields.]

CHAPTER 7

X-RAYS

AT A GLANCE			
7.1	Introduction	7.2	Production of X-rays
7.3	Spacing between Three Dimensional Lattice Planes	7.4	The Absorption of X-rays
7.5	X-ray Absorption Edges	7.6	Bragg's Law
7.7	The Bragg X-ray Spectrometer	7.8	The Powder Crystal Method
7.9	(a) The Laue Method	7.9	(b) Rotating-Crystal Method
7.10	Symmetry Operations	7.11	X-ray Spectra
7.12	Characteristic X-ray Spectrum	7.13	Moseley's Law
7.14	Compton Scattering	7.15	X-ray Crystallography
7.16	Elements of Symmetry	7.17	Bravais Lattices
7.18	Miller Indices	7.19	Typical Crystal Structures

7.1 Introduction

X-rays were discovered by Roentgen in 1895. X-rays are electromagnetic waves of short wavelengths in the range of 10 Å to 0.5Å. The longer wavelength end of the spectrum is known as the "*soft* X-rays" and the shorter wavelength end is known as "*hard* X-rays."

7.2 Production of X-rays

The Coolidge tube. X-rays are produced when fast moving electrons are suddenly stopped by a solid target. A Coolidge tube is shown in Fig. 7.1. The tube is exhausted to the best

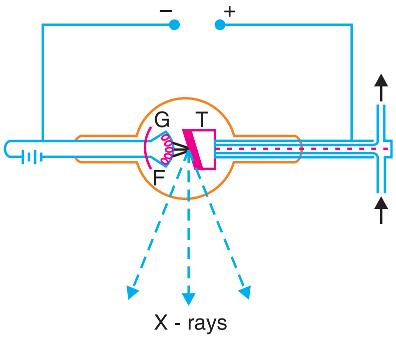
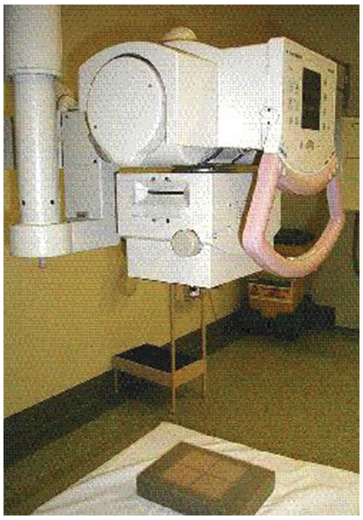


Fig. 7.1

possible vacuum of the order of 10^{-5} *mm* of mercury. The cathode consists of a tungsten filament (*F*) heated by a low tension battery. Thermionic electrons emitted by the filament are accelerated towards the target (*T*) by a high P.D. maintained between *F* and *T*. The filament is placed inside a metal cup *G* to focus the electrons on to the target. The target must be cooled to remove the heat generated in it by continuous electron-bombardment. The usual method is to mount the target material on a hollow copper tube through which cold water is continuously circulated. The target is made of a metal like tungsten or molybdenum having a high melting point and a high atomic number. Metals with high atomic number give more energetic and intense X-rays when used as targets.



X-ray Tube Assembly and Light Beam Diaphragm.

In the Coolidge tube, the intensity and frequency of X-rays can be easily controlled.

(1) The intensity of X-rays depends on the number of electrons striking the target per second. The number of electrons given out by the filament is proportional to its temperature, which can be adjusted by varying the current in the filament circuit. Therefore, the intensity of X-rays varies with the filament current.

(2) The frequency of X-rays emitted depends on the voltage between the cathode and the anode (target). Let *V* be the accelerating potential across the tube. If *e* is the charge on the electron, the work done on the electron in moving from the cathode to the anticathode = *eV*. The electron thus acquires K.E. which is converted into X-rays, when the electron strikes the target. If ν_{max} is the maximum frequency of the X-rays produced, then $h\,\nu_{max} = eV$.

$\therefore$ The minimum wavelength produced by an X-ray tube = $\lambda_{min} = \frac{c}{{}_{\text{max}}} = \frac{hc}{eV}$

7.3 Spacing between Three Dimensional Lattice Planes

Consider *O* as the origin and *OX*, *OY*, *OZ* as three rectangular axes (Fig. 7.1*a*). Let a reference plane pass through *O*, and consider a set of parallel planes defined by Miller indices (*hkl*). Beside reference plane through *O*, if next plane passes through *A*, *B* and *C* then the respective intercepts are *a/h*, *b/k* and *c/l*. If *ON* is the normal between this plane and the reference plane, then *ON* = *d* is the interplanar spacing.

Let us now find the value of *d* in terms of *a*, *b* and *c*. Let the normal make angles α, β, γ with crystal axes such that ∠ *NOX* = α, ∠ *NOY* = β, and ∠ *NOZ* = γ.

We have from Fig. 7.1*a*,

$$d = \frac{a}{h}\cos\alpha = \frac{b}{k}\cos\beta = \frac{c}{l}\cos\gamma$$

or $$\cos\alpha = \frac{d}{(a/h)},\ \cos\beta = \frac{d}{(b/k)},\ \cos\gamma = \frac{d}{(c/l)}.$$

But $$\cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1$$

$$\therefore \quad \left[\frac{d}{(a/h)}\right]^2 + \left[\frac{d}{(b/k)}\right]^2 + \left[\frac{d}{(c/l)}\right]^2 = 1$$

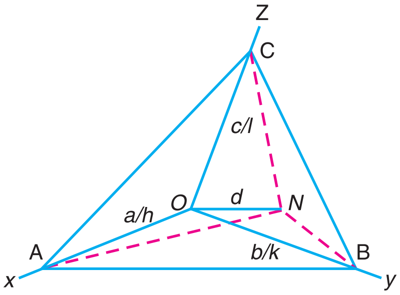


Fig. 7.1 (a)

or $$d^2\left[\frac{h^2}{a^2}+\frac{k^2}{b^2}+\frac{l^2}{c^2}\right] = 1$$

or $$d = \left[\frac{h^2}{a^2}+\frac{k^2}{b^2}+\frac{l^2}{c^2}\right]^{-½}$$

For a simple cubic lattice, $a = b = c$.

Hence for such a lattice, inter-planar separation is given by

$$d_{kkl} = \frac{a}{\sqrt{(h^2+k^2+l^2)}}$$

$$d_{100} = \frac{a}{\sqrt{(1^2+0^2+0^2)}} = a; \quad d_{110} = \frac{a}{\sqrt{(1^2+1^2+0^2)}} = \frac{a}{\sqrt{2}}$$

$$d_{111} = \frac{a}{\sqrt{(1^2+1^2+1^2)}} = \frac{a}{\sqrt{3}}$$

Thus the separation between successive (100), (110) and (111) planes are a, $a/\sqrt{2}$ and $a/\sqrt{3}$ respectively. Hence the ratio of their separation is

$$d_{100} : d_{110} : d_{111} = 1 : \frac{1}{\sqrt{2}} : \frac{1}{\sqrt{3}}$$

Similarly, for *fcc lattice*,

$$d_{100} = a/2; \; d_{110} = a/2\sqrt{2} \text{ and } d_{111} = a/\sqrt{3}$$

For *bcc lattice*

$$d_{100} = a/2; \; d_{110} = a/\sqrt{2} \text{ and } d_{111} = a/2\sqrt{3}.$$

EXAMPLE 1. *Calculate the interplanar spacing for a* (321) *plane in a simple cubic lattice whose lattice constant is* 4.2×10^{-10}*m.* **(*Rohilkhand* 93)**

SOL. $a = b = c = 4.2 \times 10^{-10}$m

and $$d_{hk1} = \frac{a}{\sqrt{(h^2+k^2+l^2)}}$$

For the plane (321), $h = 3$, $k = 2$ and $l = 1$

$$\therefore \quad d_{321} = \frac{4.2\times10^{-10}}{\sqrt{(3^2+2^2+1^2)}}\, m = \mathbf{1.1\times10^{-10}\,m}$$

EXAMPLE 2. *In a tetragonal lattice,* $a = b = 0.25$ *nm and* $c = 0.18$ *nm. Deduce the lattice spacings between* (111) *planes.* **(*Agra* 1974; *Rohilkhand* 81 S)**

SOL. The lattice spacings for a given set of parallel planes (hkl) is given by

$$d = \left[\frac{h^2}{a^2}+\frac{k^2}{b^2}+\frac{l^2}{c^2}\right]^{-½}$$

Here $h = 1$, $k = 1$ and $l = 1$

Also $a = b = 0.25$ nm and $c = 0.18$ nm

$$\therefore \quad d_{111} = \left[\frac{1^2}{(0.25)^2}+\frac{1^2}{(0.25)^2}+\frac{1^2}{(0.18)^2}\right]^{-½} = \mathbf{0.126\,nm}$$

7.4 The Absorption of X-rays

All materials through which X–rays pass absorb them to some extent. If a sheet of any substance is interposed in the path of a homogeneous beam of X–rays, its intensity decreases. Let I_0 be the intensity of the incident X–ray beam. Let I be the intensity of the beam after it has traversed a thickness dx of the absorber. Then, the decrease in intensity (dI) is observed to follow the equation $dI = -\mu I dx$. Here μ is called the linear absorption coefficient which depends on the wavelength of X-ray used and the nature of the absorbing material. μ has the dimensions of reciprocal length (L^{-1}) and its units are therefore m^{-1}.

Now, $$\frac{dI}{I} = -\mu\, dx \text{ or } I = I_0 e^{-\mu x}$$

or $$I = I_0 e^{-(\mu/\rho)\times\rho x} = I_0 e^{-\mu_m m}$$

$\mu/\rho = \mu_m$, is called the mass absorption coefficient and m is the mass of unit area of the absorbing sheet. Theory shows that the mass absorption coefficient (μ_m) varies as the cube of the incident wavelength (λ) and also the atomic number (Z) of the absorber. Thus, $\mu_m = k\lambda^3 Z^3$. This explains why materials with high atomic numbers are preferred for shielding against X-rays.

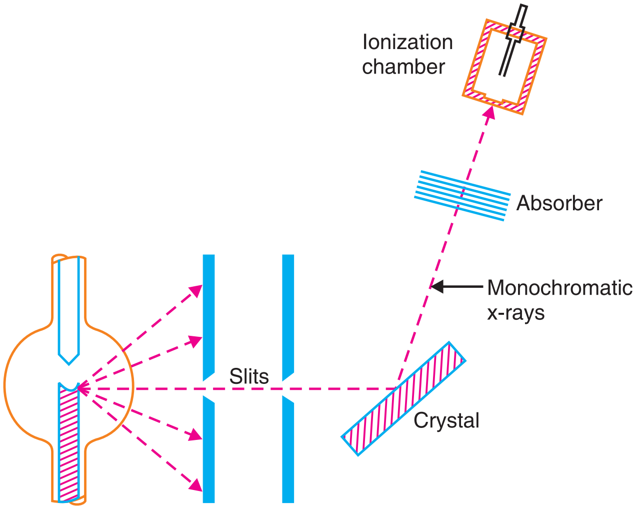


Fig. 7.2

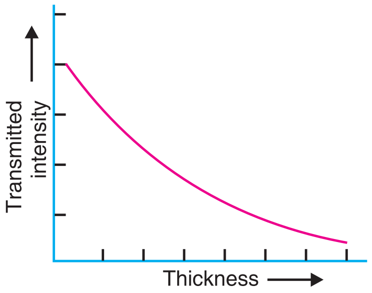
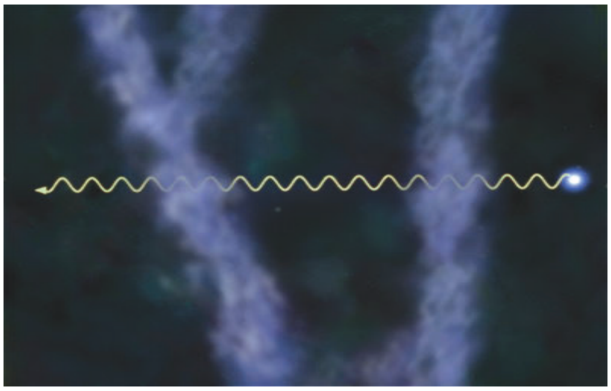


Fig. 7.3

An ionization chamber can be used to study the penetrating ability of X-rays with apparatus shown in Fig. 7.2. X–rays are collimated by slits, rendered monochromatic by a Bragg reflection, and passed through the material under study. The thickness of the absorbing material is varied and the transmitted intensity is plotted against thickness (Fig. 7.3). This is an exponential decay curve ($I = I_0 e^{-\mu x}$).



Absorption of X-rays.

7.5 X-ray Absorption Edges

The value of the linear absorption coefficient μ depends upon the X–ray wavelength and on the nature of the absorbing material. Let X-ray photons of sufficiently high energy (short wavelength) be made to fall upon a material. A major part of the incident energy of X–ray photons is spent in ejecting electrons from the K-shell of the atom. The energy of the X–ray beam is thus absorbed photo electrically, causing the ejection of the K-electrons. But if the incident radiation has a wavelength slightly longer (or the energy slightly smaller) than that required to eject the K-electrons, there is no photo-electric absorption. At this stage, there is a sudden fall in the value of the absorption coefficient (Fig. 7.4). *Such a wavelength at which there is a sudden change in the absorption coefficient of a given material for the* X-*ray beam is called* **absorption edge.**

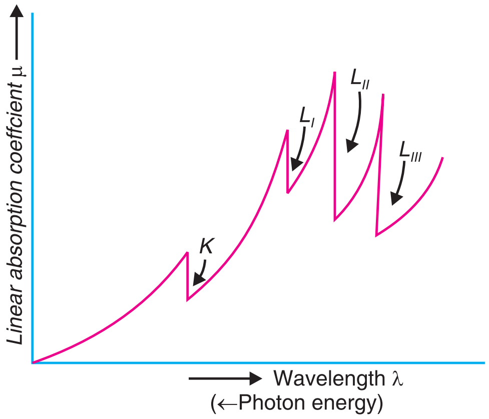
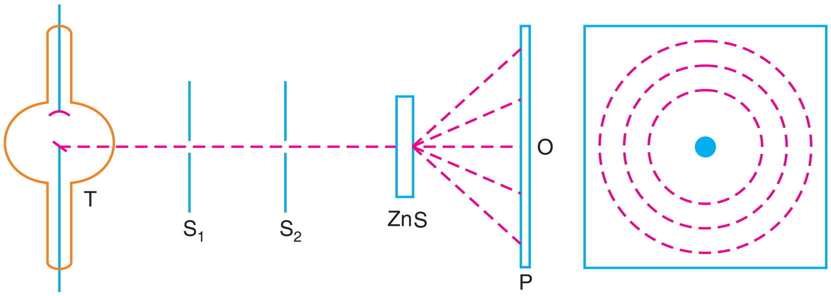


Fig. 7.4

As the wavelength is further increased, the absorption coefficient again increases. The X–ray photons now start ejecting electrons from the L-shell. There is again a sudden decrease in the value of μ at another definite wavelength $\lambda = \lambda_{LI}$. Immediately after this there are two further discontinuities at two closely lying wavelengths $\lambda_{L\,II}$ and $\lambda_{L\,III}$. In the case of ionization from the L-shell there are three absorption edges (L_I, L_{II}, L_{III}), indicating that the electrons in the L-shell can exist in three energy sub-group states.

Laue's experiment. Laue showed that if X–rays are allowed to pass through a crystal, a diffraction pattern is obtained. T is a source of X–rays. This source gives white X–rays, *i.e.*, X–ray beam of continuous range of wavelengths. A narrow beam of X–rays from the X-ray tube is collimated by two slits S_1 and S_2 and is allowed to pass through a thin crystal of zinc blende (ZnS). The transmitted beam is received on a photographic plate P [Fig. 7.5 (*i*)]. After an exposure of many hours, the plate is developed. The Laue photograph obtained on the photographic plate is shown in Fig. 7.5 (*ii*). The photograph consists of a central spot which arises due to the direct beam. The central spot is surrounded by many other fainter spots arranged in a definite pattern around O. This indicates that the incident X–ray beam has been diffracted from the various crystal planes. These spots are known as *Laue spots.* These spots are arranged according to different geometrical patterns for different crystals, depending on their structure. This experiment proved: (1) *that* X*–rays are electromagnetic waves and* (2) *that the atoms of a crystal are arranged in a regular three-dimensional lattice.*



(i) (ii)

Fig. 7.5

Bragg gave a simple interpretation of the pattern. The regular repetition of atomic, ionic or molecular units in three dimensions is called the "*lattice*". Many planes can be imagined in the lattice, containing these atoms. Some planes will be richer in atoms than others. This can be easily seen by considering two dimensional lattice. Fig. 7.5 (*iii*) represents a two-dimensional lattice, indicating lines of atom distribution. Line *AB* contains greater concentration of atoms, line *AC* less and line *AD* still less. These lines of atoms in two-dimensional lattice are analogous to planes of atoms in a three-dimensional lattice. If now a beam of X–rays is incident in the vertically downward direction, at each plane of atoms there is a partial reflection of X–rays. These reflected beams give rise to the Laue diffraction pattern. The condition to be satisfied so that X–rays reflected from a set of lattice planes may produce an intense spot in the diffraction pattern may be derived. This is known as Bragg's law.

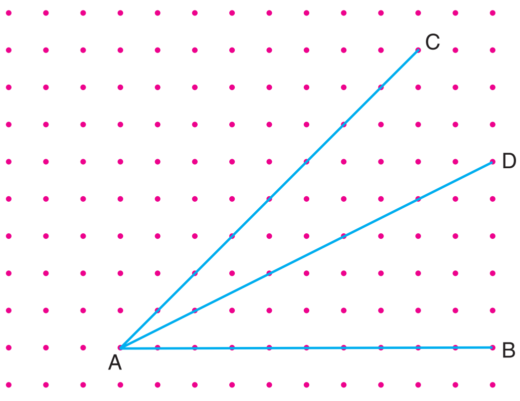


Fig. 7.5 (iii)

7.6 Bragg's Law

When monochromatic X–rays impinge upon the atoms in a crystal lattice, each atom acts as a source of scattering radiation of the same wavelength. The crystal acts as a series of parallel reflecting planes. The intensity of the reflected beam at certain angles will be maximum when the path difference between two reflected waves from two different planes is an integral multiple of λ.

Derivation of Bragg's law. Consider a set of parallel planes of atom points at a spacing d between two successive planes. Let a narrow monochromatic X–ray beam of wavelength λ be incident on the first plane at a glancing angle θ (Fig. 7.6). Consider the ray *PQ* incident on the first plane. The corresponding reflected ray *QR* must also be inclined at the same angle θ to the plane. Since X–rays are much more penetrating than ordinary light, there is only partial reflection at each plane. The complete absorption takes place only after penetrating several layers. Consider two parallel rays *PQR* and *P′Q′R′* in the beam, which are reflected by two atoms *Q* and *Q′*. *Q′* is vertically below *Q*. The ray *P′Q′R′* has a longer path than the ray *PQR*. To compute the path-difference between the two rays, from *Q* draw normals *QT* and *QS* on *P′Q′* and *Q′R′* respectively. Then the path-difference = $TQ' + Q'S = d \sin\theta + d \sin\theta = 2d \sin\theta$.

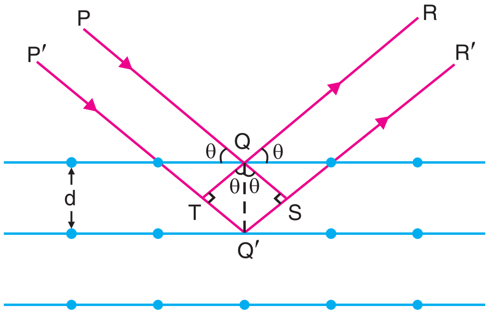


Fig. 7.6

Hence the two rays will reinforce each other and produce maximum intensity, if

$$2d \sin\theta = n\lambda,$$

where $n = 1, 2, 3$...... The integer n gives the order of the scattered beam, λ is the wavelength of the X-rays used. This equation is called *Bragg's law.*

Bragg X-ray Spectrometer.

7.7 The Bragg X-ray Spectrometer

The essential parts of a Bragg spectrometer are shown in Fig. 7.7. It is similar in construction to an optical spectrometer. It consists of three parts. (1) a source of X-rays (2) a crystal held on a circular table which is graduated and provided with vernier and (3) a detector (ionisation chamber). X–rays from an X–ray tube, limited by two narrow lead slits S_1 and S_2, are allowed to fall upon the crystal *C*. The crystal is mounted on the circular table *T*, which can rotate about a vertical axis and its position can be determined by the vernier *V*. The table is provided with a radial arm (*R*) which carries an ionisation chamber (*I*). This arm also can be rotated about the same vertical axis as the crystal. The position of this arm can be determined by a second vernier (not shown in the figure). The ionisation chamber is connected to an electrometer (*E*) to measure the ionisation current. Hence we can measure the intensity of the diffracted beam of X–rays, diffracted in the direction of the ionisation chamber. S_3 is a lead slit, to limit the width of the diffracted beam. In practice, the crystal table is geared to the ionisation chamber so that the chamber turns through 2θ when the crystal is turned through θ.

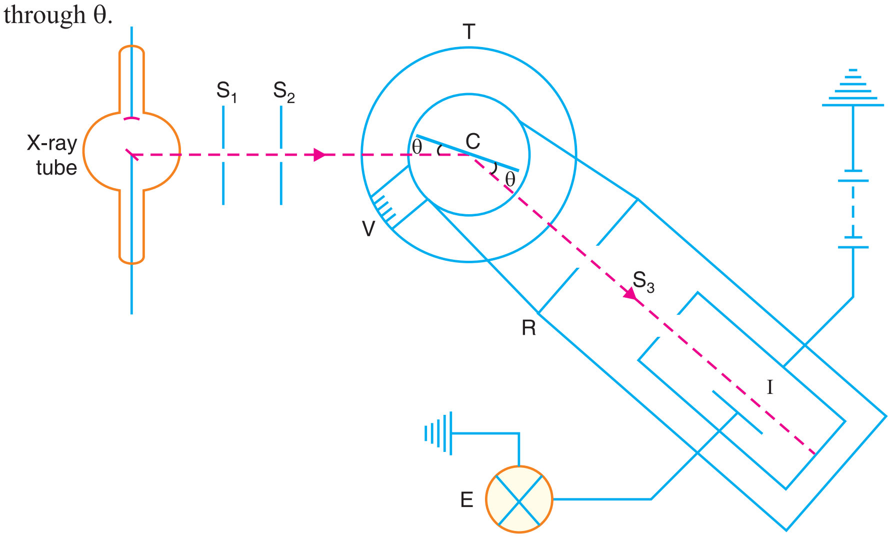


Fig. 7.7

Working. To begin with, the glancing angle θ for the incident beam is kept very small. The ionisation chamber is adjusted to receive the reflected beam till the rate of deflection is maximum. The glancing angle (θ) and the intensity of the diffracted beam (*I*) are measured. The glancing angle is next increased in equal steps, by rotating the crystal table. The ionisation current is noted for different glancing angles. The graph of ionisation current against glancing angle is drawn. The graph obtained is as in Fig. 7.8 and is called an ***X-ray spectrum.*** The prominent peaks A_1, A_2, A_3 refer to X-rays of wavelength λ. The glancing angles $\theta_1, \theta_2, \theta_3$ corresponding to the peaks A_1, A_2, A_3 are obtained from the graph. It is found that $\sin\theta_1 : \sin\theta_2 : \sin\theta_3 = 1 : 2 : 3$. This shows that A_1, A_2, A_3, refer to the first, second and third order reflections of the same wavelength. B_1, B_2, B_3 are such peaks for the first, second and third order for another wavelength (λ_2). Thus Bragg experimentally verified the relation

$$2d \sin\theta = n\lambda.$$

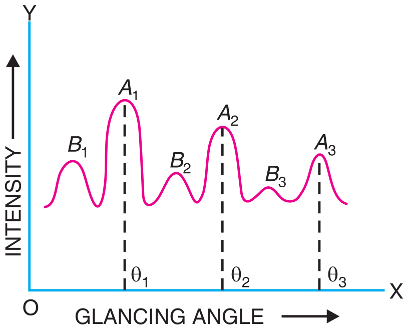


Fig. 7.8

The wavelength of X-rays is determined by using the equation $2d \sin\theta = n\lambda$. The glancing angle θ is experimentally determined as explained already for a known order. If *d* is known, λ can be calculated.

Calculation of *d*. Rocksalt (*NaCl*) possesses a cubic structure with sodium ions and chlorine ions situated alternately at corners of a cube. If *d* is the distance between two neighbouring ions and ρ is the density of the crystal, then mass of the unit cube = ρd^3. Now each corner ion is shared by 8 neighbour cubes. This is illustrated by the ion indicated by an asterisk in Fig. 7.9. Hence each ion contributes only $\frac{1}{8}$th of its mass to the cube. The unit cube is made up of 4 sodium ions and 4 chlorine ions. Hence, the mass of each cube = the mass of $\frac{1}{2}$ *NaCl* molecule. If *M* is the molecular weight of *NaCl* and N_A the Avogadro's No.,

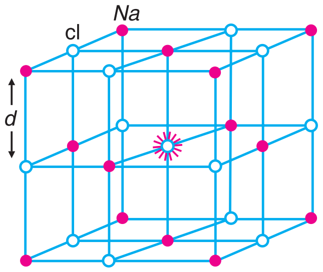


Fig. 7.9

Mass of $\frac{1}{2}$ *NaCl* molecule = $M/2N_A$

$$\therefore \qquad \rho d^3 = \frac{M}{2N_A} \text{ or } d = \sqrt[3]{\frac{M}{2N_A\,\rho}}$$

Knowing M, N_A and ρ, d can be found.

For *NaCl*, $M = 58.45$ kg ; $\rho = 2170$ kg m^{-3} and $N_A = 6.06 \times 10^{26}$.

$$\therefore \qquad d = \sqrt[3]{\frac{58.45}{2 \times (6.06 \times 10^{26}) \times 2170}} = 2.81 \times 10^{-10} \text{ m.}$$

Example 1. *The spacing between principal planes of NaCl crystal is* 2.82 Å. *It is found that first order Bragg reflection occurs at an angle of* 10°. *What is the wavelength of X–rays* ?

Sol. By Bragg equation, $2d \sin \theta = n\lambda$.

Here, $d = 2.82 \times 10^{-10}$m; $n = 1$ and $\theta = 10°$. λ = ?

$$\lambda = \frac{2d \sin \theta}{n} = \frac{2 \times (2.82 \times 10^{-10}) \sin 10°}{1} = 0.98 \times 10^{-10} \text{ m.}$$

Example 2. *Bragg's spectrometer is set for the first order reflection to be received by the detector at glancing angle of* 9º 18*'.Calculate the angle through which the detector is rotated to receive the second order reflection from the same face of the crystal.*

Sol. Let λ be the wavelength of X–rays. Let θ_1 and θ_2 be the glancing angles for the first and second orders. Then,

$$2d \sin \theta_1 = 1\lambda \text{ and } 2d \sin \theta_2 = 2\lambda.$$

$\therefore$ $\sin \theta_2 = 2 \sin \theta_1$, or $\sin \theta_2 = 2 \sin 9^0 18' = 0.3232$ or $\theta_2 = 18°48'$

Now, $\theta_2 - \theta_1 = 9º\,30'$. Hence the crystal is rotated through an angle 9°30′. The angle through which the detector is rotated is twice the angle through which the crystal is rotated. Therefore, the angle through which the detector is rotated from first order to second order $= 2\,(\theta_2 — \theta_1) = 19º$.

7.8 The Powder Crystal Method

The Laue's and Bragg's techniques for the investigation of crystal structures can be applied only if single crystals of reasonable size are available. But, in general, large crystals, without fault are difficult to obtain. Therefore, Debye and Scherrer adopted a different technique. The specimen was taken in the form of a well powdered sample of the crystal in a thin glass capsule.

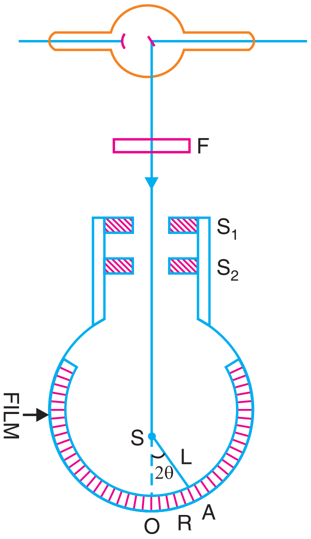


Fig. 7.10

The experimental arrangement is shown in Fig. 7.10. The X–rays from the source are made approximately monochromatic by the filter *F.* A narrow beam of these monochromatic X-rays, suitably collimated by two lead slits S_1 and S_2, falls on the powdered specimen *S*. The specimen is suspended vertically on the axis of a cylindrical camera. The photographic film is mounted round the inner surface of the camera, covering nearly the whole circumference in order to receive the beams diffracted upto 180°.

The powder specimen of the crystal can be imagined to be a collection of random oriented tiny crystals, presenting all values of glancing angles to the incident beam. For a given wavelength and a given value of *d*, there can be only one value of θ (the glancing angle) which satisfies the equation $2d \sin \theta = n\lambda$, where $n = 1$. Such reflected beams emerge out from the specimen in all directions inclined at an angle 2θ with the direction of the incident beam, because millions of tiny crystals in the specimen are random oriented. The reflected rays will be on the surface of a cone, vertex at the specimen, base on the photographic film and having a semi-vertical angle 2θ. The traces obtained on the photographic film will be as in Fig. 7.11.

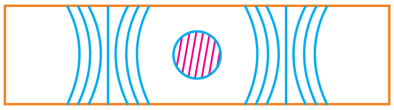


Fig. 7.11

Let *L* be the radius of the cylindrical camera. The direct beam strikes the film at *O*. Suppose a spectrum with glancing angle θ is found at *A* which is at a distance *R* from *O*. Then $\theta = R / 2L$. Using this value of θ in the Bragg's equation and knowing the value of λ, *d* (spacing of the plane involved) can be calculated.

The powder method has been employed in the study of microcrystalline substances like metals, alloys, carbons, fluorescent powders and other forms where single crystals are not available.

7.9 (a) The Laue Method

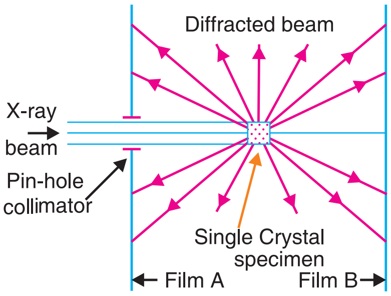


Fig. 7.12 (a)

A single crystal is held stationary in a beam of X-rays of continuous wavelength. The crystal selects and diffracts the discrete values of λ for which planes exist of spacing *d* and incidence angle θ satisfying the Bragg law. The experimental arrangement is shown in Fig. 7.12 (*a*). A source is used that produces a beam of X-rays over a wide range of wavelengths preferably from 0.02 nm to 0.2 nm. A pinhole arrangement produces a well-collimated beam. The dimensions of the crystal are usually less than 1 mm. Flat films are placed in front of and behind the specimen. Since λ covers a continuous range, the crystal selects that particular wavelength which satisfies Bragg's law at the present orientation. A diffracted beam emerges at the corresponding angle. The diffracted beam is then recorded as a spot on the film. But since the wavelength corresponding to a spot is not measured, one cannot determine the actual values of the interplanar spacings—only their ratios. Therefore one can determine the shape but not the absolute size of the unit cell.

The pattern will show the symmetry of the crystal : if a crystal has a fourfold axis of symmetry parallel to the beam, the Laue pattern will show fourfold symmetry. The Laue method is widely used to orient crystals for solid state experiments.

This method is, however, not suitable for determining the crystal structure. This is because, out of a continuous range of wavelengths, several wavelengths are reflected in different orders from a single plane, so that different orders of reflection may overlap on a single spot. This makes difficult the measurement of the reflection intensity of individual spots and hence of the missing reflections. The missing reflections would have been an important step in the determination of the crystal structure. However, the shape of the unit cell can be established from the symmetry of the pattern.

7.9 (b) Rotating-Crystal Method

In this method, a single crystal is rotated about a fixed axis (usually vertical), in a beam of monoenergetic X-rays. The variation in the angle θ brings different atomic planes into position for reflection. The experimental arrangement is shown in Fig. 7.12 (*b*).

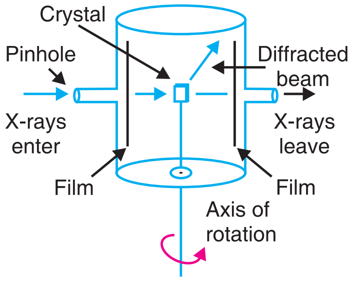


Fig. 7.12 (b)

The crystal is usually about 1 mm in diameter and is mounted on a spindle which can be rotated. A photographic film is placed on the inner side of a cylinder concentric with the axis of rotation. The incident X-ray beam is monochromatised by a filter or by reflection from an earlier crystal. The incident beam is diffracted from a given crystal plane whenever in course of rotation the value of θ satisfies the Bragg equation. Beams from all planes parallel to the vertical rotation axis will lie in the horizontal plane. Planes with other orientations will reflect in layers above and below the horizontal plane. However, there will be no reflected beams from a plane which always contains the incident beam during the whole rotation and from ones whose spacing is so small that $\lambda/2\,d > 1$. The reflected spots on the film form parallel lines. Since λ is known, the spacing d may be calculated from $2d \sin\theta = n\,\lambda$.

7.10 Symmetry Operations

A symmetry operation is one which leaves the crystal invariant. Symmetry operations may be grouped into three classes:

1. Translation operations
2. Point operations 3. Hybrid operations

1. Translation Operation. A translation operation is defined as the displacement of a crystal parallel to itself by a crystal translation vector defined by $\boldsymbol{T} = n_1\boldsymbol{a} + n_2\boldsymbol{b}$.

2. Point Operations

***(i)* The mirror reflection:** In this operation, the reflection of a structure at a mirror plane *m* passing through a lattice point leaves the crystal unchanged. Fig. 7.13 (*a*) shows two mirror planes in a two-dimensional crystal.

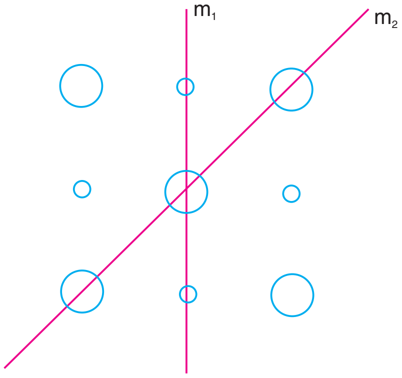


Fig. 7.13 (a)

***(ii)* Inversion:** A crystal structure has an inversion symmetry if for every lattice point of position vector $\boldsymbol{r}$, there is a corresponding lattice point at the position $-\boldsymbol{r}$ [Fig. 7.13 (*b*)]. The origin about which the position vector $\boldsymbol{r}$ is chosen is called the *centre of inversion* and is represented by *i*.

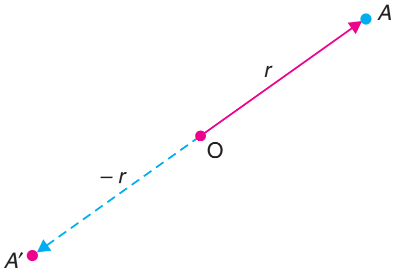


Fig. 7.13 (b)

***(iii)* Rotation:** A structure is said to possess rotational symmetry about an axis if rotation of the structure about this axis by an angle ϕ gives an unchanged configuration of the original structure. The angle ϕ which satisfies this condition is given by

$$\phi = \frac{2\pi}{n};\ n = 1, 2, 3, \ldots\ldots\ldots$$

where the rotation axis is n-fold. In a crystal lattice only one, two, three, four, and six-fold rotation axes of symmetry are possible. Fig. 7.13 (*c*) shows a 6-fold axis of rotation.

3. Hybrid Operations

(*i*) Rotoreflection: It is the combination of an n-fold rotation followed by a reflection in a plane perpendicular to the rotation axis. It is denoted by n/m, where $n = 1, 2, 3, 4, 6$.

(*ii*) Rotoinversion: This is the combination of an n-fold rotation followed by an inversion and is denoted by $\overline{n}$ with $n = 1, 2, 3, 4, 6$.

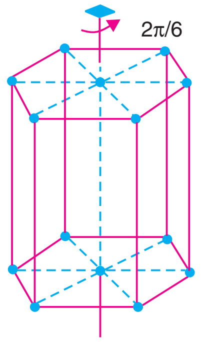


Fig. 7.13 (c)

The axis of rotation in rotoreflection and rotoinversion is referred to as improper axis.

(*iii*) Screw translation: In this the n-fold rotation axis is coupled with the translation parallel to rotation axis. The new symmetry element is called a *screw axis* and is denoted by n_m.

(*iv*) Glide reflection: We can combine a reflection with a translation parallel to the reflection plane to achieve congruence in some crystals. This combination is called a glide reflection. The translation component of a glide reflection is expressed as the combination of translations by fractions of repeat distances in a crystal. For example, the translation in a diagonal glide may be written as

$$\boldsymbol{t} = \frac{1}{2}\boldsymbol{a}+\frac{1}{2}\boldsymbol{b},\ \frac{1}{2}\boldsymbol{b}+\frac{1}{2}\boldsymbol{c} \ \text{ or } \ \frac{1}{2}\boldsymbol{c}+\frac{1}{2}\boldsymbol{a}.$$

The list of symmetry operations is not without a very special symmetry operation, known as the *identity operation.* This operation means doing no operation on the crystal or keeping the crystal undisturbed. Thus every crystal has the identity operation. It is generally denoted by the symbol I.

Close-packed Crystal Structures. There are two ways of arranging equivalent spheres in a regular array to minimise the interstitial volume. One has hexagonal symmetry and is called the hexagonal close-packed structure. The other has cubic symmetry and is called the face-centred cubic structure.

Spheres may be arranged in a single closest-packed layer by placing each sphere in contact with six others (Fig. 7.14 *a*). The centres of the spheres are at points marked A. A second and identical layer of spheres can be placed over this, with centres over the points marked B (or equivalently C). If the second layer goes in over B, there are two nonequivalent choices for a third layer. It can go in over A or over C. If the third layer goes in over A, the sequence is *ABABAB*.... and the structure is hcp (Fig. 7.14 *b*). Examples of elements which have this structure are Cd, Mg, Ti, Zn, etc.

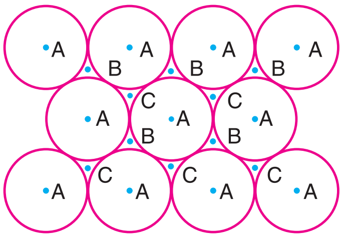


Fig. 7.14 (a)

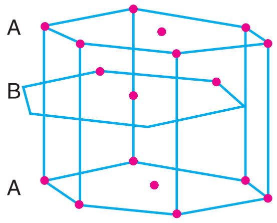


Fig. 7.14 (b)

If the third layer goes in over C, the sequence is *ABCABCABC* and the structure is fcc (Fig. 7.14 *c*). Elements having fcc structure are Cu, Ag, Au, Al, Pd, Pt, etc.

In either structure, each atom has 12 nearest neighbours. So the *coordination number* is 12. The packing fraction is 0.74 for both the fcc and hcp structures.

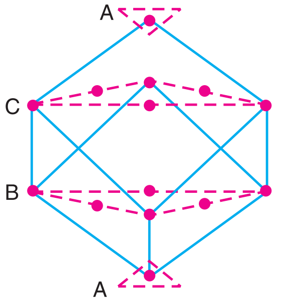


Fig. 7.14 (c)

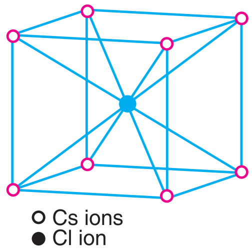


Fig. 7.14 (d)

The Cesium Chloride Structure. The structure of CsCl is shown in Fig 7.14*d*. The Cs^+ ions form an sc lattice interlocked with another sc lattice formed by the Cl^- ions. There is one molecule per primitive cell with ions at the corners 000 and body centred positions $\frac{1}{2}\frac{1}{2}\frac{1}{2}$ of the simple cubic space lattice. Each ion is at the centre of a cube of ions of the opposite kind, so that the coordination number is 8. TlBr, TlI, NH_4Cl, CuPd, CuZn (β-brass), AgMg, LiHg, AlNi, BeCu are representative crystals having the CsCl arrangement.

7.11 X-ray Spectra

Urey and his co-workers analysed the X-ray beam emitted from an X-ray tube using different potential differences and same target. Using tungsten as the target and different potential differences, the intensities of the rays produced are plotted against wavelength. Fig. 7.15 represents these graphs. For applied *P.D.*s 30 *kV*, 40 *kV* and 50 *kV*, the spectrum is *white*. But, for the applied P.D. 70 *kV*, two sharp peaks are seen. The sharp peaks show the *line* or *characteristic radiation*. The line spectra are absent, till the P.D. is greater than a particular value. The smoothly varying curves represent the *continuous spectrum.* The superimposed lines on the continuous background constitute the *characteristic spectrum.*

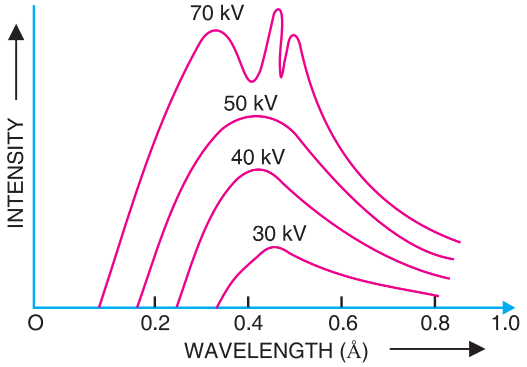


Fig. 7.15

Main features of continuous X-ray spectrum

(1) For each anode potential, there is a minimum wavelength (λ_{min}) below which no radiation is emitted. Above this critical value, the intensity of the radiation increases rapidly with increasing wavelengths and after reaching a maximum, decreases gradually. The intensity never reaches zero showing that the radiation contains all possible wavelengths above the minimum limit.

(2) When the voltage across the X-ray tube is increased, λ_{min} is shifted towards smaller values. Duane and Hunt

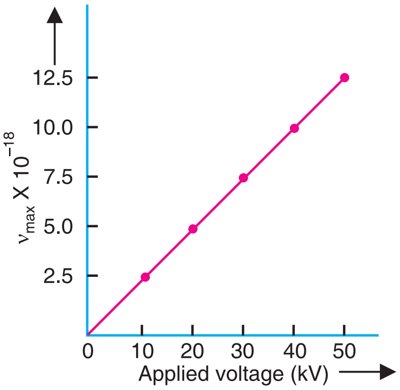


Fig. 7.16

showed that λ_{min} is inversely proportional to the applied voltage V or ν_{max} is directly proportional to V. If the limiting frequencies (ν_{max}) are plotted against the applied voltages (V), a straight line graph passing through the origin is obtained (Fig. 7.16). This empirical law of Duane and Hunt is expressed analytically as

$$eV = h\nu_{\max} = \frac{hc}{\lambda_{\min}} \cdot$$

Explanation. eV is the K.E. of the bombarding electron. If the entire K.E. of the electron striking the target is converted into the energy of the X-ray photon, then, $eV = h\nu_{max}$ according to Einstein's theory.

But $$\nu_{max} = \frac{c}{\lambda_{\min}}$$

$$\therefore \qquad eV = \frac{hc}{\lambda_{\min}} \quad \text{or} \quad \lambda_{\min} = \frac{hc}{eV} \quad \text{or} \quad \lambda_{\min} \propto \frac{1}{V} \cdot$$

Most of the electrons that generate X-ray photons give up only a part of their energy in this way. Therefore, most of the X-radiation is of longer wavelength than λ_{min}. Thus, the continuous spectrum is the result of the inverse photoelectric effect, with electron kinetic energy (eV) being transformed into photon energy ($h\nu$).

7.12 Characteristic X-ray Spectrum

There are two methods of producing characteristic X-rays.

(1) The characteristic X-rays of an element can be excited by using the element as the target in the X-ray tube and thus subjecting it to direct bombardment by electrons. For each target there is a minimum potential below which the line spectra do not appear. This critical P.D. below which the line spectra do not appear, is different for different targets. Molybdenum shows up the line spectra only if the P.D. is above 35 kV.

(2) Characteristic X-rays of an element can also be excited by allowing primary X-rays from a hard X-ray tube to fall on the element. The primary X-rays must be harder than the characteristic X-rays to be produced. The peaks obtained in the X-ray spectrum (Fig. 7.15) give us the line spectrum which is characteristic of the element used in the target. The group of lines of shortest wavelength is called the K-series. Usually two lines of this series are detected. These lines are termed as K_α and K_β lines in the order of decreasing wavelengths. The next group is called the L-series of longer wavelengths (L_α, L_β, L_γ, etc.). For heavier elements a third series, called the M-series has been detected.

Origin of characteristic X-rays. This can be understood in terms of Bohr's theory. Suppose an atom in the target of an X-ray tube is bombarded by a high-speed electron and a K-electron is removed. A vacancy is created in the K-shell. This vacancy can be filled up by an electron from either of L, M or N shells or a free electron. These possible transitions can result in the K_α, K_β lines and the limiting line. Similarly, the longer wavelength L-series originates when an L electron is knocked out of the atom, the M-series when an M electron is knocked out and so on (Fig. 7.17).

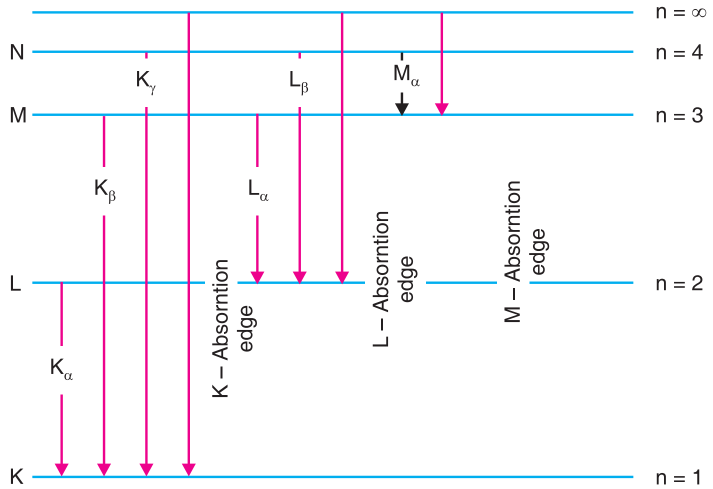


Fig. 7.17

Note

It is clear that continuous spectra and line spectra both emitted by the same target are of different origin. The continuous spectrum is the result of the inverse photoelectric effect, with electron K.E. being transformed into photon energy hv. The line spectrum has its origin in electronic transitions within atoms that have been disturbed by the incident electrons.

EXAMPLE 1. *An X-ray tube with Cu target is operated at 28 kV. The smallest glancing angle for NaCl crystal for the CuK$_\alpha$ line is 15.8°. Find the wavelength of this line. Also find the glancing angle for photons at the short wavelength limit.* (*d for NaCl* = 0.282 *nm*).

SOL. From Bragg's relation, $2d \sin\theta = n\lambda$. Here, $n = 1$

$$\therefore \qquad \lambda = 2d\sin\theta = 2 \times (0.282 \text{ nm}) \times \sin 15.8° = 0.153 \text{ nm}$$

$$\lambda_{min} = \frac{hc}{eV} = \frac{6.62 \times 10^{-34} \times 3 \times 10^{8}}{1.6 \times 10^{-19} \times 28 \times 10^{3}} \text{m} = 0.0496 \text{nm}$$

Let θ' be the glancing angle corresponding to λ_{min}. Then

$$\theta' = \sin^{-1}\left(\frac{\lambda_{\min}}{2d}\right) = \sin^{-1}\left(\frac{0.0496 \text{ nm}}{2 \times 0.282 \text{ nm}}\right) = \mathbf{5.1°}$$

EXAMPLE 2. *X-rays are more penetrative than visible light. Why?*

SOL. Wavelength of X-rays is very small compared to the wavelength of visible light *i.e.*, energy of X-rays is very high compared to the energy of visible light. Hence X-rays are more penetrative than visible light.

7.13 Moseley's Law

Moseley plotted the square root of the frequencies ($\sqrt{\nu}$) of a given line (say K_α) against the atomic numbers (Z) of the elements emitting that line. Moseley obtained a straight line as shown in Fig. 7.18. The same linear relation was found to hold good for any line in any series. He concluded, therefore, that atomic number (and not atomic weight) is the fundamental property of elements.

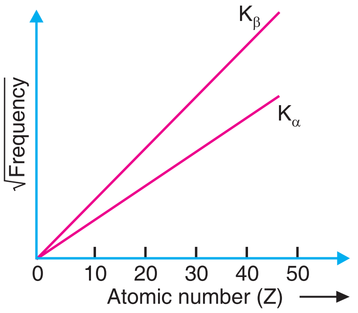


Fig. 7.18

Moseley's law: statement. *The frequency of a spectral line in X-ray spectrum, varies as the square of the atomic number of the element emitting it,* or $\nu \propto Z^2$. Moseley's law may be written as $\sqrt{\nu} = a\,(Z - b)$. Here, Z is the atomic number of the element and a and b are constants depending upon the particular line.

Explanation, according to Bohr's theory. Bohr's theory of hydrogen spectrum gives the frequency of a spectral line as $\nu = Z^2 Rc \left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right)$ where R is Rydberg's constant and c the velocity of light. Considering the K_α line, we can regard it as originating from the transition of electron from second to first orbit. Now, $n_1 = 1$ and $n_2 = 2$.

So frequency of k_α line $= \nu = Z^2 Rc \left(\frac{1}{1^2} - \frac{1}{2^2}\right) = \frac{3}{4} cRZ^2$

This approximately corresponds to Moseley's law.

Importance of Moseley's law. (1) According to this law, it is the atomic number and not atomic weight of an element which determines its characteristic properties, both physical and chemical. Therefore, the atoms must be arranged in the periodic table according to their atomic numbers and not according to their atomic weights. This would remove some discrepancies in the order of certain elements from the point of view of their atomic weights. For example, argon $_{18}\text{Ar}^{40}$ comes before potassium ($_{19}\text{K}^{39}$), cobalt ($_{27}\text{Co}^{58.9}$) comes before nickel ($_{28}\text{Ni}^{58.7}$), etc. So the arrangement is correct in the order of atomic number.

(2) Moseley's work has also helped to perfect the periodic table by (*i*) the discovery of new elements, *e.g.*, hafnium (72), illinium (61), masurium (43), rhenium (75), etc., and (*ii*) the determination of the atomic numbers of rare-earths and fixing their positions in the periodic table.

Example. *The wavelength of the* L_α *X-ray line of plantinum* (*atomic number* 78) *is* 1.321 Å. *An unknown substance emits* L_α *X-rays of wavelength* 4.174 Å. *Calculate the atomic number of the unknown substance. Given* $b = 7.4$ *for* L_α *lines.*

Sol. Moseley's law is

$$\sqrt{\nu} = a\,(Z - b)$$

Let ν_1 and ν_2 be the frequencies of the L_α line of platinum and the unknown substance respectively. Let their atomic numbers be Z_1 and Z_2 respectively. Then

$$\sqrt{\nu_1} = a(Z_1 - b) \text{ and } \sqrt{\nu_2} = a\,(Z_2 - b)$$

or

$$\sqrt{\frac{\nu_1}{\nu_2}} = \frac{Z_1 - b}{Z_2 - b} \quad \text{or} \quad Z_2 - b = (Z_1 - b)\sqrt{\frac{\nu_2}{\nu_1}}$$

Since $c = \nu\lambda$, we have $\nu_2/\nu_1 = \lambda_1/\lambda_2$

$$\therefore \quad Z_2 - b = (Z_1 - b)\sqrt{\frac{\lambda_1}{\lambda_2}} \quad \text{or } Z_2 = b + (Z_1 - b)\sqrt{\frac{\lambda_1}{\lambda_2}}$$

Here, $b = 7.4$; $Z_1 = 78$; $\lambda_1 = 1.321$ Å and $\lambda_2 = 4.174$ Å

$$\therefore \quad Z_2 = 7.4 + (78 - 7.4)\sqrt{1.321/4.174} = 47.12$$

The atomic number of the unknown substance is 47.12.

7.14 Compton Scattering

The Compton Effect. *Compton discovered that when X-rays of a sharply defined frequency were incident on a material of low atomic number like carbon, they suffered a change of frequency* *on scattering. The scattered beam contains two wavelengths. In addition to the expected incident wavelength, there exists a line of longer wavelength. The change of wavelength is due to loss of energy of the incident X-rays. This elastic interaction is known as Compton effect.*

Compton Scattering Experiment.

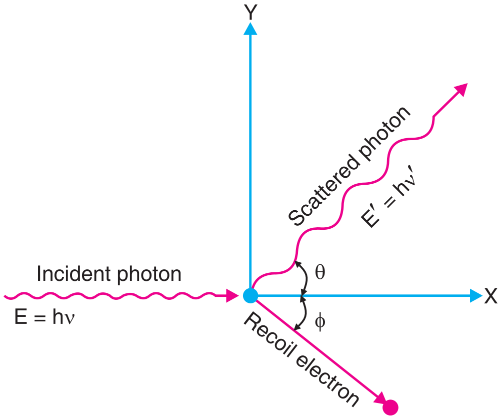


Fig. 7.19

In the case of incoherent scattering, a scattered beam undergoes not only deviation in its direction but also change of wavelength occurs. In Compton effect, there is a change in wavelength of the scattered beam along with the change in its direction. Hence Compton effect is an incoherent scattering.

This effect was explained by Compton on the basis of quantum theory of radiation. The whole process is treated as a particle collision event between X-ray photon and a loosely bound electron of the scatterer. In this process, both momentum and energy are conserved. In the photon-electron collision, a portion of the energy of the photon is transferred to the electron. As a result, the X-ray proceeds with less than the original energy (and therefore has a lower frequency or a higher wavelength).

The incident photon with an enegy $h\nu$ and momentum $h\nu/c$ strikes an electron at rest. The initial momentum of the electron is zero and its initial energy is only the rest mass energy, m_0c^2. The scattered photon of energy $h\nu$' and momentum $h\nu'/c$ moves off in a direction inclined at an angle θ to the original direction. The electron acquires a momentum mv and moves at an angle ϕ to the original direction. The energy of the recoil electron is mc^2 (Fig. 7.19).

According to the principle of conservation of energy,

$$h\nu + m_0c^2 = h\nu' + mc^2 \qquad \text{...(1)}$$

Considering the x and y components of the momentum and applying the principle of conservation of momentum,

$$\frac{h\nu}{c} = \frac{h\nu'}{c}\cos\theta + mv\cos\phi \qquad \text{...(2)}$$

and

$$0 = \frac{h\nu'}{c}\sin\theta - mv\sin\phi \qquad \text{...(3)}$$

From (2), $mvc\cos\phi = h(\nu - \nu'\cos\theta)$ ...(4)

From (3), $mvc\sin\phi = h\nu'\sin\theta$ ...(5)

Squaring and adding (4) and (5),

$$\begin{aligned} m^2v^2c^2 &= h^2(\nu^2 - 2\nu\nu'\cos\theta + \nu'^2\cos^2\theta) + h^2\nu'^2\sin^2\theta \\ &= h^2(\nu^2 - 2\nu\nu'\cos\theta) + h^2\nu'^2 = h^2(\nu^2 - 2\nu\nu'\cos\theta + \nu'^2) \qquad \text{...(6)} \end{aligned}$$

From (1),

$$mc^2 = h(\nu - \nu') + m_0c^2$$

$$\therefore \quad m^2c^4 = h^2(\nu^2 - 2\nu\nu' + \nu'^2) + 2h(\nu - \nu')m_0c^2 + m_0^2c^4 \qquad \text{...(7)}$$

Subtracting (6) from (7),

$$m^2c^2\,(c^2 - v^2) = -\,2h^2\nu\nu'(1-\cos\theta)+2h(\nu-\nu')\,m_0c^2+m_0^2c^4 \qquad ...(8)$$

The value of $m^2c^2\,(c^2 - v^2)$ can be obtained from the relativistic formula

$$m = \frac{m_0}{\sqrt{(1-v^2/c^2)}}.\ \text{Squaring,}$$

$$m^2 = \frac{m_0^2}{1-v^2/c^2} = \frac{m_0^2\,c^2}{c^2-v^2}$$

$$\therefore \qquad m^2c^2\,(c^2 - v^2) = m_0^2c^4 \qquad ...(9)$$

From (8) and (9),

$$m_0^2c^4 = -2h^2\,\nu\nu'(1-\cos\theta)+2h(\nu-\nu')\,m_0\,c^2+m_0^2\,c^4$$

$$\therefore \qquad 2h(\nu-\nu')\,m_0\,c^2 = 2h^2\nu\nu'(1-\cos\theta)$$

or $$\frac{\nu-\nu'}{\nu\nu'} = \frac{h}{m_0\,c^2}(1-\cos\theta) \ \text{ or } \ \frac{1}{\nu'}-\frac{1}{\nu} = \frac{h}{m_0\,c^2}(1-\cos\theta)$$

or $$\frac{c}{\nu'}-\frac{c}{\nu} = \frac{h}{m_0\,c}(1-\cos\theta)$$

or $$\lambda'-\lambda = \frac{h}{m_0\,c}(1-\cos\theta) \qquad ...(10)$$

$\therefore$ The change in wavelength $= d\lambda = \dfrac{h}{m_0c}(1-\cos\theta)$

This relation shows that $d\lambda$ is independent of the wavelength of the incident radiations as well as the nature of the scattering substance. $d\lambda$ depends upon the angle of scattering only.

Case 1. When $\theta = 0$, $\cos\theta = 1$ and hence $d\lambda = 0$

Case 2. When $\theta = 90°$, $\cos\theta = 0$ and hence

$$\mathrm{d}\lambda = \frac{h}{m_0c} = \frac{6.63\times10^{-34}}{(9.11\times10^{-31})\times(3\times10^{8})}\,m = 0.0243\ \text{Å}$$

This is known as *Compton wavelength.*

Case 3. When $\theta = 180°$, $\cos\theta = -1$ and hence $d\lambda = 2h/m_0c = 0.0485$ Å. $d\lambda$ has the maximum value at $\theta = 180°$.

Experimental verification. Monochromatic X-rays of wavelength λ are allowed to fall on a scattering material like a small block of carbon (Fig. 7.20). The scattered X-rays are received by a Bragg spectrometer and their wavelength is determined. The spectrometer can freely swing in an arc about the scatterer. The wavelength of the scattered X-rays is measured for different values of the scattering angle. The experimental results obtained by Compton are shown in Fig. 7.21. In the scattered radiation in addition to the incident wavelength (λ), there exists a line of longer wavelength (λ'). The "*Compton shift*"$d\lambda$ is found to vary with the angle at which the scattered rays are observed.

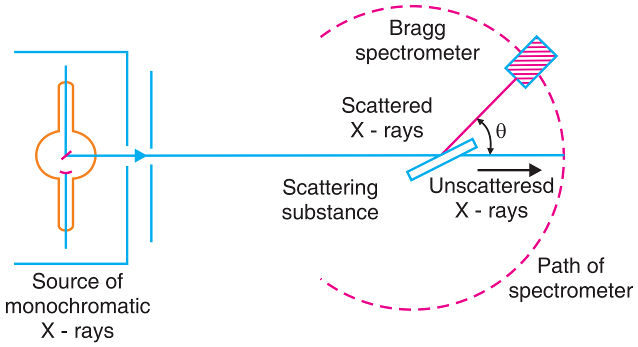


Fig. 7.20

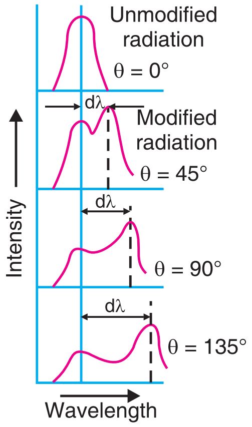


Fig. 7.21

Direction of Recoil electron. Dividing Eq. (5) by Eq. (4), we get

$$\tan\phi = \frac{h\nu'\sin\theta}{h(\nu-\nu'\cos\theta)} = \frac{\nu'\sin\theta}{(\nu-\nu'\cos\theta)}. \qquad ...(12)$$

Using Eq. (10), we get

$$\frac{1}{\nu'} = \frac{1}{\nu}+\frac{h}{m_0\,c^2}(1-\cos\,\theta) = \frac{1}{\nu}+\frac{h}{m_0\,c^2}.\,2\sin^2\frac{\theta}{2}$$

or

$$\nu' = \frac{\nu}{1+\left(\dfrac{h\nu}{m_0\,c^2}\right)2\sin^2\dfrac{\theta}{2}} = \frac{\nu}{1+2\beta\,\sin^2\left(\dfrac{\theta}{2}\right)} \text{ where } \beta = \frac{h\nu}{m_0\,c^2} \qquad ...(13)$$

Substituting this value of ν′ in Eq. (12), we get

$$\tan\phi = \frac{\nu\sin\theta\Big/\left[1+2\beta\sin^2\left(\dfrac{\theta}{2}\right)\right]}{\left[\nu-\left\{\nu\cos\theta\Big/\left(1+2\beta\sin^2\dfrac{\theta}{2}\right)\right\}\right]} = \frac{\cot\left(\dfrac{\theta}{2}\right)}{(1+\beta)}$$

∴

$$\tan\phi = \frac{\cot\left(\dfrac{\theta}{2}\right)}{1+\left(\dfrac{h\nu}{m_0\,c^2}\right)} \qquad ...(14)$$

Kinetic Energy of Recoil electron. The K.E. of recoil electron is the difference between the energies of incident and scattered photons, *i.e.*,

$$\text{K.E.} = h\nu - h\nu'.$$

$$\text{K.E.} = h\nu - h\left[\frac{\nu}{1+2\beta\sin^2(\theta/2)}\right] = h\nu\left[\frac{2\beta\sin^2(\theta/2)}{1+2\beta\sin^2(\theta/2)}\right] \qquad ...\ (15)$$

where $\beta = h\nu / m_0c^2$

EXAMPLE 1. *X-rays of wavelength* 0.7080 *Å are scattered from a carbon block through an angle of* 90° *and are analysed with a calcite crystal, the interplanar distance of whose reflecting planes is* 3.13 *Å. Determine the angular separation, in the first order, between the modified and the unmodified rays.*

Sol. Wavelength of the modified rays = $\lambda' = \lambda + \frac{h}{m_0 c}(1 - \cos\theta)$

$$= 0.7080 \times 10^{-10}\,\text{m} + \left(\frac{6.63 \times 10^{-34}}{9.11 \times 10^{-31} \times 3 \times 10^8}\right)\text{m} = 0.7323\,\text{Å}$$

Let θ and θ' be the angles of Bragg reflections corresponding to the wavelengths λ and λ'. Then, for $n = 1$ (first order),

$$2d \sin\theta = n\lambda = 0.7080 \times 10^{-10}\ \text{m}$$

and $$2d \sin\theta' = n\lambda' = 0.7323 \times 10^{-10}\ \text{m}.$$

Here, $$d = 3.13 \times 10^{-10}\ \text{m};\ \therefore\ \theta = 6°30' \text{ and } \theta' = 6°\,43'.$$

$\therefore$ The angular separation, in the first order, between the modified and unmodified rays $\Big\}$ $= \theta' - \theta = 13'$

7.15 X-ray Crystallography

Types of Solids. Solids are of two types : crystalline and amorphous. A substance is said to be crystalline when the arrangement of the units of matter inside it is regular and periodic; the units of matter may be ions, atoms or molecules. Crystals are usually bounded by plane faces. Though two crystals of the same substance may look different in external appearance, the angles between corresponding faces are always the same. Crystals can be distinguished from amorphous substances by their *anisotropy*, *i.e.*, the dependence of quantities like refractive index on the direction in which they are measured in the crystal. In amorphous-solids, there is no order in the arrangement of their constituent particles.

Periodic Arrays of Atoms. An ideal crystal is constructed by the infinite regular repetition in space of identical structural units in the form of parallelepipeds. In the simplest crystals such as copper, silver, gold and the alkali metals the structural unit contains a single atom. More generally the structural unit contains several atoms or molecules. A crystal may be composed of more than one chemical element (as in *Na Cl*) or it may contain associated groups of identical atoms (as in H_2). We describe the structure of all crystals in terms of a single periodic *lattice,* but with a group of atoms attached to each lattice point. This group of atoms is called the *basis*; the basis is repeated in space to form the crystal.

Crystal Translation vectors and Lattices. An ideal crystal is composed of atoms arranged on a lattice defined by three *fundamental translation vectors* **a, b, c** such that the atomic arrangement looks the same in every respect when viewed from any point **r** as when viewed from the point

$$\boldsymbol{r'} = \boldsymbol{r} + n_1\boldsymbol{a} + n_2\boldsymbol{b} + n_3\boldsymbol{c}, \qquad \text{...(1)}$$

where n_1, n_2, n_3 are arbitrary integers. The set of points r' specified by (1) for all values of the integers n_1, n_2, n_3 defines a *lattice.* The lattice and the translation vectors **a, b, c** are said to be *primitive* if *any* two points ***r, r'*** from which the atomic arrangement looks the same always satisfy (1) with a suitable choice of the integers n_1, n_2, n_3.

We often use the primitive translation vectors to define the *crystal axes* ***a***, ***b***, ***c*** although nonprimitive crystal axes may be used when they are more convenient or simpler. The crystal axes ***a***, ***b***, ***c*** form three adjacent edges of a parallelepiped. If there are lattice points only at the corners of the parallelepiped, then it is a *primitive parallelepiped.*

A *lattice translation operation* or *crystal translation operation* is defined as the displacement of a crystal parallel to itself by a crystal translation vector

$$\boldsymbol{T} = n_1\boldsymbol{a} + n_2\boldsymbol{b} + n_3\boldsymbol{c} \qquad \text{...(2)}$$

A vector ***T*** connects any two lattice points.

Illustration. Fig. 7.22 gives the diagram of a two dimensional lattice with lattice vectors $\boldsymbol{a}$ and $\boldsymbol{b}$. The atomic arrangement in the crystal looks exactly the same to an observer at $\boldsymbol{r}'$ as to an observer at $\boldsymbol{r}$, provided that the vector $\boldsymbol{T}$ which connects $\boldsymbol{r}'$ and $\boldsymbol{r}$ may be expressed as an integral multiple of the vectors $\boldsymbol{a}$ and $\boldsymbol{b}$. In this illustration $\boldsymbol{T} = -\boldsymbol{a} + 3\boldsymbol{b}$. The vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ are primitive translation vectors of the two-dimensional lattice.

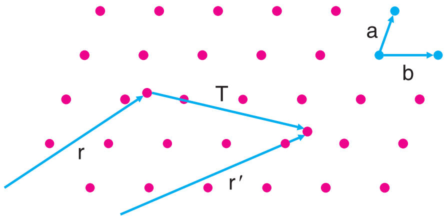


Fig. 7.22

The crystal lattice. A crystal is constructed by the infinite repetition in space of identical structural units (atoms, molecules or ions). One can replace each unit by a geometrical point. The result is a pattern of points having the same geometrical properties as the crystal. This geometrical pattern is the *crystal lattice* or simply the *lattice.* The points are called lattice points.

Definition. *The regular pattern of points which describe the three dimensional arrangement of particles* (*atoms, molecules or ions*) *in a crystal structure is called the crystal lattice or space lattice.*

Fig. 7.23 (*a*) represents an array of points in two dimensions. Note that environments about any two points is the same. Hence, Fig. 7.23 (*a*) represents a lattice.

Basis. A crystal structure is formed by associating with every lattice point a unit assembly of atoms or molecules identical in composition. This unit assembly is called the basis. Fig. 7.23 (*b*) shows a particular arrangement of two different ions which represents the basis. The number of atoms in the basis may be as low as one, as for many metals but there are structures for which the basis exceeds 1000 atoms.

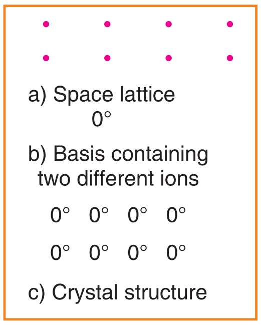


Fig. 7.23

The crystal structure. A crystal structure is formed by the addition of a basis to every lattice point *i.e.*, *Lattice + Basis = Crystal structure.* The crystal structure is real whereas the lattice is imaginary. Fig. 7.23 clearly illustrates the difference between the crystal lattice, the basis, and the crystal structure. The crystal structure is formed by the addition of the basis (*b*) to every lattice point of the lattice (*a*). By looking at (*c*) we can recognise the basis and then we can abstract the space lattice.

Unit cell. The atoms in a crystal are arranged in a periodic array. It is therefore possible to isolate a representative *unit cell* in each variety of crystal. This unit cell may be a group of ions, atoms or molecules. Now, we can construct the crystal by repeatedly translating the unit cell in three dimensions.

Definition. *The smallest portion of a space lattice which can generate the complete crystal by repeating its own dimensions in various directions is called unit cell.*

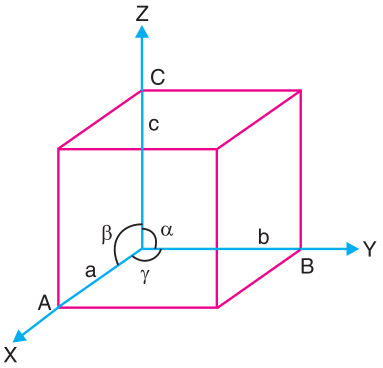


Fig. 7.24

A unit cell (Fig. 7.24) is defined by the length of its edges and by the angles between them. $OA = a$, $OB = b$ and $OC = c$ are the dimensions of the unit cell. The angles between a, b; b, c and c, a are denoted by γ, α and β respectively. They are called interfacial angles.

The vectors **a, b, c** define the axes of the crystal.

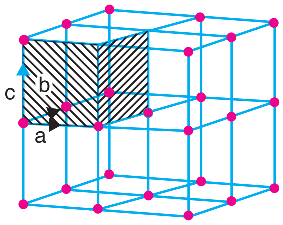


Fig. 7.25

Primitive Lattice Cell. The parallelepiped (Fig. 7.25) defined by primitive axes $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ is called a *primitive cell.* A primitive cell is a type of *unit cell.* A unit cell will fill all space under the action of suitable crystal translation operations; a primitive cell is a minimum-volume unit cell. There is a density of one lattice point per primitive cell. There are lattice points at the eight corners of the parallelepiped, but each corner point is shared among the eight cells which touch there. The volume V_c of a primitive cell defined by primitive axes $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ is

$$V_c = |(\boldsymbol{a} \times \boldsymbol{b}) \cdot \boldsymbol{c}|$$

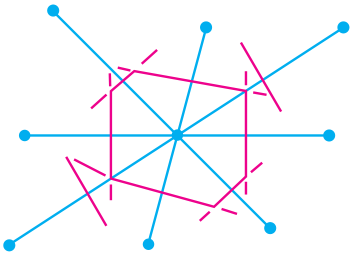


Fig. 7.26

Another way of choosing a cell of equal volume V_c is shown in Fig. 7.26.

(*i*) Draw lines to connect a given lattice point to all nearby lattice points.

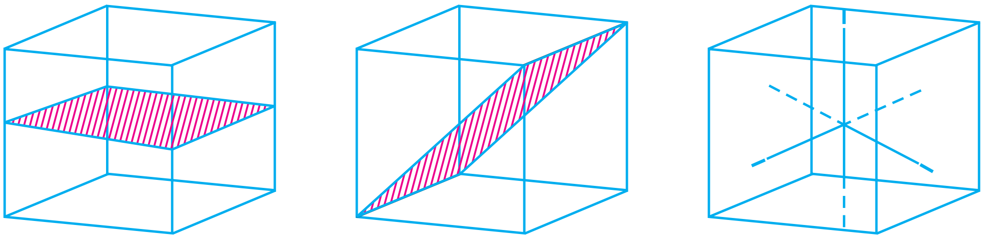
(*ii*) At the midpoint and normal to these lines, draw new lines or planes. The smallest volume enclosed in this way is the *Wigner-Seitz primitive cell*. The basis associated with a lattice point of a primitive cell may be called a primitive basis. No basis contains fewer atoms than a primitive basis contains.

7.16 Elements of Symmetry

What is meant by a symmetry ? Formally, we might say that a symmetry of a particular kind exists, when a certain operation leaves something unchanged. A candle is symmetric about a vertical axis because it can be rotated about that axis without changing in appearance or any other feature.

Crystals possess different symmetries. The symmetry operations of a crystal carry the crystal structure into itself. Symmetry operations performed about a point or line are called point-group symmetry. The point-group symmetry elements possessed by a crystal are :

(1) **Plane of symmetry.** A crystal is said to have a *plane of symmetry*, when it is divided by an imaginary plane into two halves, such that one is the mirror image of the other. For example, a cube has nine planes of symmetry. Fig. 7.27 (*a*) represents a plane of symmetry parallel to the faces of a cube. There are three such planes. Fig. 7.27 (*b*) represents one of the six diagonal planes of symmetry.



(*a*) (*b*) (*c*)

Fig. 7.27

(2) **Axis of symmetry.** This is an axis such that, if the crystal is rotated around it through some angle, the crystal remains invariant. The axis is called *n*-fold if the angle of rotation is 360/*n*.

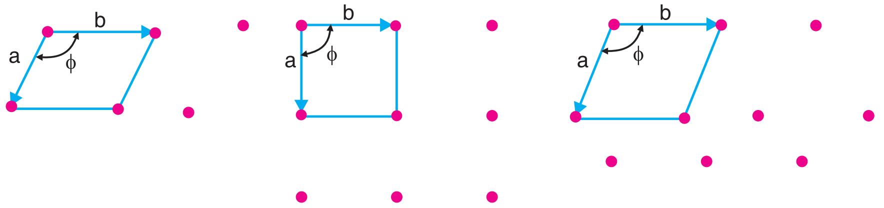
Example. If equivalent configuration occurs after rotation of 180°, 120° and 90°, the axes of rotation are known as two-fold, three-fold and four-fold, axes of symmetry respectively. For example, Fig. 7.27 (*c*) shows the three tetrad (four-fold) axes of a cube.

(3) **Centre of symmetry.** It is a point such that any line drawn through it will meet the surface of the crystal at equal distances on either side. It is important to mention here that a crystal may possess a number of planes or axes of symmetry but it can have only one centre of symmetry. *Example.* A cube has a single centre of symmetry at the centre of the cube.

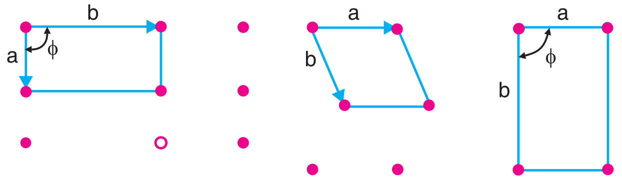
Point Group. *A lattice point group is defined as the collection of the symmetry operations which, when applied about a lattice point, leave the lattice invariant.* 32 such groups are possible. Every known crystal can be assigned to one of these 32 crystal classes.

Space Groups. When identical objects are placed at the lattice points of a *space lattice* a regular special array is obtained, which has, perhaps, *symmetry elements* over and above those of the *point group* of the original lattice. Thus, because of some symmetry property of the identical objects (*i.e.*, of the *unit cells* in a crystal), there may exist *glide* (= reflection + translation) *planes* and *screw* (= rotation + translation) *axes* which define operations turning the structure into itself. It has been shown that there are only 230 different types of symmetry possible for such a system. Thus 230 three dimensional space groups or "structures" exist in total.

Two-Dimensional Latice Types. There are five Bravais lattices in two-dimensions (Fig. 7.28). These are listed in Table 7.1.



(*a*) Oblique lattice $|a| \neq |b|$; $\phi \neq 90°$ (*b*) Square lattice $|a| = |b|$; $\phi = 90°$ (*c*) Hexagonal lattice $|a| = |b|$; $\phi = 120°$



(*d*) Rectangular lattice $|a| \neq |b|$; $\phi \neq 90°$ (*e*) Centred rectangular lattice; axes are shown for both the primitive cell and for the rectangular unit cell, for which $|a| \neq |b|$; $\phi = 90°$

Fig. 7.28

TABLE 7.1. The five two-dimensional lattice types
(The notation mm means that two mirror lines are present)

Lattice	*Conventional unit cell*	*Axes of conventional unit cell*	*Point-group symmetry of lattice about lattice points*
Oblique	Parallelogram	$a \neq b$ $\phi \neq 90°$	2
Square	Square	$a = b$ $\phi = 90°$	4 *mm*
Hexagonal	60° rhombus	$a = b$ $\phi = 120°$	6 *mm*
Primitive rectangular	Rectangle	$a \neq b$ $\phi = 90°$	2 *mm*
Centred rectangular	Rectangle	$a \neq b$ $\phi = 90°$	2 *mm*

7.17 Bravais lattices

Bravais showed that only 14 types of space lattice are possible. Fig. 7.29 shows the 14 *Bravais lattices*. From the figure it is clear that seven sets of axes are sufficient to construct the 14 Bravais lattices. This leads to the classification of all crystals into seven crystal systems:

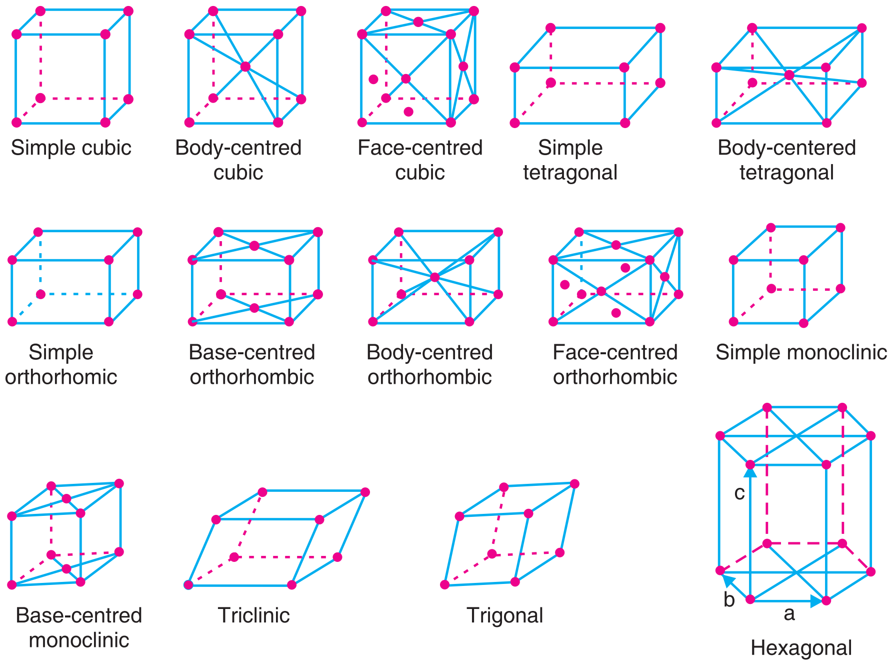

Simple cubic
Body-centred cubic
Face-centred cubic
Simple tetragonal
Body-centered tetragonal
Simple orthorhomic
Base-centred orthorhombic
Body-centred orthorhombic
Face-centred orthorhombic
Simple monoclinic
Base-centred monoclinic
Triclinic
Trigonal
c
b
a
Hexagonal

Fig. 7.29

(1) Cubic (2) Tetragonal (3) Ortho-rhombic (4) Monoclinic (5) Triclinic (6) Trigonal (sometimes called rhombohedral) and (7) hexagonal.

Table 7.2 gives the seven crystal systems, divided into fourteen Bravais lattices. Note that a simple lattice has points only at the corners, a body-centred lattice has one additional point at the centre of the cell, and a face-centred lattice has six additional points, one on each face.

TABLE 7.2

The seven crystal systems, divided into fourteen Bravais lattices

System	*Bravais lattice*	*Unit cell characteristics*	*Characteristic symmetry elements*
Cubic	Simple Body-centred Face-centred	$a = b = c$ $\alpha = \beta = \gamma = 90°$	Four 3-fold rotation axes (along cube diagonal)
Tetragonal	Simple Body-centred	$a = b \neq c$ $\alpha = \beta = \gamma = 90°$	One 4-fold rotation axis
Orthorhombic	Simple Base-centred Body-centred Face-centred	$a \neq b \neq c$ $\alpha = \beta = \gamma \neq 90°$	Three mutually orthogonal 2-fold rotation axes
Monoclinic	Simple Base-centred	$a \neq b \neq c$ $\alpha = \beta = 90° \neq \gamma$	One 2-fold rotation axis

Triclinic	Simple	$a \neq b \neq c$ $\alpha \neq \beta \neq \gamma \neq 90°$	None
Trigonal (rhombohedral)	Simple	$a = b = c$ $\alpha = \beta = \gamma \neq 90°$	One 3-fold rotation axis
Hexagonal	Simple	$a = b \neq c$ $\alpha = \beta = 90°$ $\gamma = 120°$	One 3-fold rotation axis

There is an interesting point about cubic lattices. This is called the coordination number. *It is defined as the number of particles 'immediately' adjacent to each particle in the crystal lattice.* It is characteristic of a given space lattice and is determined by an inspection of the model. In a simple cubic lattice, each particle is adjoined by six other particles and so the coordination number is six. The coordination number of body-centred and face-centred cubic lattice are 8 and 12 respectively.

7.18 Miller Indices

Miller introduced a system to designate a plane in a crystal. He introduced a set of three numbers to specify a plane in a crystal. This set of three numbers is known as Miller indices of the concerned plane.

Procedure for finding Miller indices. (1) Determine the intercepts of the plane along the axes *X*, *Y* and *Z* in terms of the lattice constants *a*, *b* and *c*.

(2) Determine the reciprocals of these numbers.

(3) Find the least common denominator (*lcd*) and multiply each by this *lcd*.

The result is written in the form (*hkl*) and is called the Miller indices of the plane.

Example. *Let a plane have intercepts 4, 1 and 2 on the three axes. The reciprocals are 1/4, 1 and 1/2. lcd is 4. Multiplying each by 4, we get 1, 4 and 2. Hence (142) are the Miller indices of the plane.*

Notes

1. For an intercept at infinity, the corresponding index is zero.
2. If a plane cuts an axis on the negative side of the origin, corresponding index is negative.
3. The indices (*hkl*) do not define a particular plane but a set of parallel planes. Thus the planes whose intercepts are 1,1,1; 2, 2, 2; –3, –3, –3, etc., are all represented by the same set of Miller indices.
   Miller indices of some important planes in a cubic crystal are shown in Fig. 7.30.

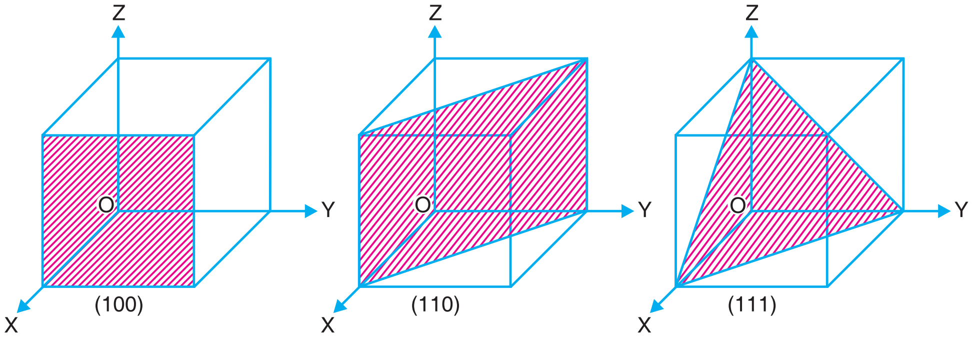


Fig. 7.30

4. It is only the ratio of the indices which is important in this notation. The (622) planes are the same as (311) planes.

EXAMPLE. *In a crystal, a lattice plane cuts intercepts of 2a, 3b and 6c along the axes, where a, b and c are primitive vectors of the unit cell. Determine the Miller indices of the given plane.*

SOL. If the given plane cuts intercepts *ra, sb, tc* along the three axes, we have

$$ra : sb : tc = 2a : 3b : 6c$$

where a, b and c are primitive vectors of the unit cell.

$$r : s : t = 2 : 3 : 6$$

$$\frac{1}{r} : \frac{1}{s} : \frac{1}{t} = \frac{1}{2} : \frac{1}{3} : \frac{1}{6} = 3 : 2 : 1$$

∴ Miller indices of the plane are (321).

7.19 Typical Crystal Structures

1. Fluorite structure. Fig. 7.31 (*a*) shows the atomic arrangement in fluorite structure. The fluorite structure (CaF_2) is cubic, with Ca atoms at cube corners and face centers and with *F* atoms at all quarter-way position along the cube diagonals $\left(\frac{1}{4}\frac{1}{4}\frac{1}{4}, \frac{1}{4}\frac{3}{4}\frac{3}{4} \text{ etc.}\right)$. These typically ionic structures are found frequently among alloys of metals with elements of group IV A, V A and VI A.

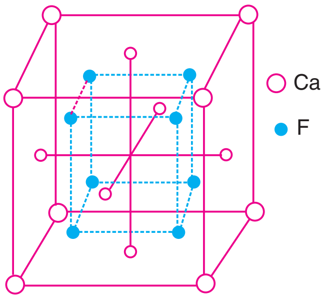


Fig. 7.31 (a)

2. Perovskite Structure. Fig. 7.31 (*b*) shows the atomic arrangement in perovskite structure. Perovskite with general formula $A\,B\,X_3$ (A and B are cations and X anions) has a unit cell of cubic symmetry. It consists of a framework of corner linked octahedra with anions at the corners, the B cations at the centres of the octahedra and the A cation at the centre of the space between the octahedra. Examples are $BaTiO_3$ (ferroelectric), $SrTiO_3$, $CaTiO_3$ (mineral), superconductors like Ba ($Pb_{0.8}$ $Bi_{0.2}$) O_3, ($Ba_{0.6}$ $K_{0.4}$) BiO_3.

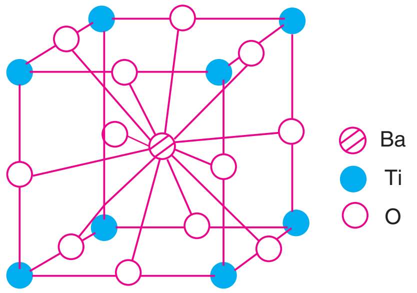


Fig. 7.31 (b)

3. Crystals of Alkali Metals. Crystals of alkali metals (Li, Na, K, Rb, Cs) are typical representatives of the body centred cubic (BCC) structure. The unit cell is non-primitive with two lattice points and the basis of one atom. The positions of atoms at the two non-equivalent lattice points are 000 and $\frac{1}{2}\frac{1}{2}\frac{1}{2}$. The coordination number is eight.

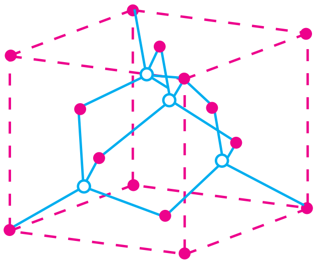


Fig. 7.32

Structure of zinc blende (*ZnS*). The zinc blende structure consists of two *fcc* sub-lattices. *Zn* atoms occupy the lattice points of one sub-lattice and *S* atoms of the other (Fig.7.32). Thus each *Zn* atom has four *S* neighbours at the corners of a regular tetrahedron and *vice versa.* In this case, there are double bonds between atoms which are partially covalent and partially ionic. Many semiconductor compounds crystallize in the Zinc-blende structure.

Determination of crystal structure by Bragg's law. To find the structure of a crystal, the *lattice constant* is determined by using different planes of the crystal as reflecting surface for the same known wavelength of X-rays. The lattice constants are expressed as ratios with one of them taken as unity. This ratio comes to be different for different crystals. The experimentally observed ratio is then compared with the calculated ratio and thus a particular structure may be identified.

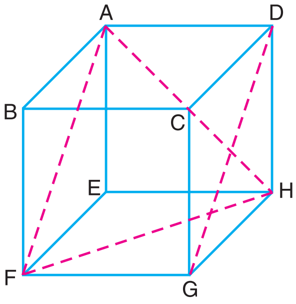


Fig. 7.33

Structure of *KCl* crystal. Consider a crystal of simple cubic system in which an atom lies at each corner of the cube (Fig. 7.33). The structure is repeated in all directions constituting a *space lattice.* In this case, there are three sets of planes which are rich in atoms.

(*i*) The first set of planes such as *ABFE*, *CDHG*, *ADHE*, *BCGF*, *ABCD* and *EFGH* are all alike. They are called (100) planes. Let the distance between the consecutive planes be d_1.

(*ii*) The second set of planes consist of parallel planes like *ADGF* or *ABGH* (110 planes) inclined at 45° to the cubic faces. If d_2 is the spacing between these planes, then $d_2 = d_1/\sqrt{2}$.

(*iii*) The third set of planes consist of planes like *AFH* (111 planes). The separation between these planes (d_3) can also be calculated in terms of d_1. ($d_3 = d_1/\sqrt{3}$)

$\therefore$ For a simple cubic system, $d_1 : d_2 : d_3 = 1 : \dfrac{1}{\sqrt{2}} : \dfrac{1}{\sqrt{3}}$.

The first order spectrum from the (100), (110) and (111) planes of *KCl* was observed at the glancing angles 5.22°, 7.30° and 9.05° respectively.

Then since $d \propto 1/\sin\theta$,

$$d_1 : d_2 : d_3 = \frac{1}{\sin 5.22^\circ} : \frac{1}{\sin 7.30^\circ} : \frac{1}{\sin 9.05^\circ} = 1 : 1/\sqrt{2} : 1/\sqrt{3}.$$

This is in agreement with the theoretical values. Hence *KCl* belongs to the simple cubic system. But, from these studies, no information is obtained as to the nature of the constituents at the lattice corners, whether they are molecules of *KCl* or ions of *K* and *Cl* alternately. The answer to this question is obtained from observations made on *NaCl* crystal.

Structure of *NaCl* crystal. The results obtained with *NaCl* crystal for the same three planes are identical with the previous case. However, a weak reflection at the (111) face at about 5° is also obtained. Taking this into account, we find that $d_1 : d_2 : d_3 = 1 : 1/\sqrt{2} : 2/\sqrt{3}$, which agrees with the calculated values for a face-centred cube. Hence *NaCl* appears to belong to the face-centred cubic system.

The difference between the *KCl* and *NaCl* crystals is explained by the fact that "the number of electrons (atomic number) in the atom determines the amount by which an atom scatters X-rays". The atomic numbers of potassium (19) and chlorine (17) are not very much different. If we assume that all the atoms are identical, it means that the face-centred arrangement of *NaCl* (Fig. 7.34) becomes a simple cubic arrangement for *KCl.* But, in the case of *NaCl* crystal, there is difference between the atomic numbers of *Na* (11) and *Cl* (17). It

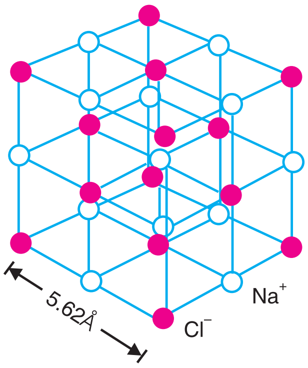


Fig. 7.34

means that their scattering powers are different and hence give rise to a face-centred cubic lattice. From the above discussion, we can say that "both *NaCl* and *KCl* have the same structure".

EXERCISE

1. Give an account of the production, properties and applications of X-rays. What factors influence (*i*) cut off wavelength (*ii*) the wavelengths of lines emitted from an X-ray tube?

2. Describe Laue's experiment and point out its significance.

3. Derive Bragg's law for X-ray diffraction in crystals. How is it verified ? Describe and explain the X-ray spectrometer method of determining wavelength of X-rays.

4. Explain how X-ray spectrometer may be used to study the structure of crystals.

5. Describe the powdered crystal method of studying crystal structure.

6. Distinguish between continuous X-radiations and characteristic X-ray emission spectra of the element and explain the origin of the latter on Rutherford-Bohr model.

7. Describe Moseley's work on X-rays. What is Moseley's law ? What is its importance ?

8. Give the theory of Compton effect and briefly explain its experimental verification.

9. Explain the terms : (*i*) (*a*) space lattice, (*b*) cleavage plane and (*c*) lattice constant.

(*ii*) Explain Miller indices. Draw the planes with the following indices in a cubic unit cell : (200), (110), (111), (010).

10. The P.D. across an X-ray tube is 10^5 volt. What is the maximum frequency of X-rays emitted? What is the corresponding wavelength ? [**Ans.** 2.4×10^{19} Hz; 1.25×10^{-11} ***m***]

11. What should be the P.D. through which an electron must be accelerated from rest, so that when it falls on a target the short wavelength limit of its continuous spectrum shall be 10^{-10}m? Also estimate the maximum speed of the electron. [**Ans.** 12.4 *kV*; $6.61 \times 10^7\ ms^{-1}$]

12. Calculate the heat generated per minute in the target of an X-ray tube if the voltage applied is 20 *kV* and the current is 10 *mA*. How much power is radiated if the efficiency is 0.2% ? [**Ans.** 12000 *J*; 0.4 *W*]

13. An X-ray tube operates at 12 kV. Find the maximum speed of the electrons striking the anticathode. Mass of electron = 9.1×10^{-31} *kg*; charge of electron = 1.602×10^{-19}*C*.

$$\left[\textbf{Hint. } \frac{1}{2} mv^2_{\max} = eV. \ \therefore \ v_{\max} = \sqrt{\frac{2eV}{m}} = 6.5\times10^7 \ ms^{-1}\right]$$

14. Find the smallest glancing angle at which K_α radiation of molybdenum of wavelength 0.7 Å., will be reflected from calcite crystal of spacing 3.036 Å. At what angle will be third-order reflection ?

[**Ans.** 6°37′ ; 20°18′]

15. *NaCl* has its principal planes spaced at 2.820 Å. The first order of Bragg reflection is located at 10°. Calculate (*a*) the wavelength of the *X*-rays and (*b*) the angle for the second order Bragg reflection.

[**Ans.** 0.9790 Å.; 20° 19′]

16. An X-ray tube is operated at 0.04 Megavolt. The radiations from the tube are analysed by a Bragg spectrometer. Calculate (*a*) shortest wavelength limit of the spectrum and (*b*) the glancing angle for the 2nd order of this wavelength, given the crystal grating spacing = 3.036 Å. [**Ans.** 0.309 Å.; 2°55′]

17. In an X-ray diffraction by the powder method, the radius of the cylindrical photographic film is 8 cm while the radius of a first order spectrum line on the photograph is 5 cm. If the spacing of the reflecting atomic planes is 2.75 Å., calculate the wavelength of the X-rays used.

[**Hint.** $\tan 2\theta = \frac{R}{L} = \frac{5}{8}$.

$\therefore\ 2\theta = 32°$ or $\theta = 16°$. $n\lambda = 2d \sin \theta$.

But $n = 1$.

$\therefore\ \lambda = 2 \times 2.75 \times \sin 16°$ Å. = 1.52 Å.]

18. The K_α line from molybdenum has a wavelength of 0.7078 Å. Calculate the wavelength of the K_α line of copper. Atomic number of molybdenum = 42. Atomic number of copper = 29.

[**Hint.** From Moseley's law, $(Z–1)^2 \lambda$ = a constant. Let λ' be the wavelength of K_α line of copper. Then, $(42–1)^2 \times 0.7078 = (29–1)^2 \lambda'$ $\therefore$ λ' =1.517 Å.]

19. Monochromatic X-radiation of wavelength 0.124 Å., undergoes Compton effect from a carbon block. Calculate the wavelength that is scattered through 180°. [**Ans.** 0.1724 Å.]

20. Photon of energy 1.02 MeV undergoes Compton scattering through 180°. Calculate the energy of the scattered photon.

[**Hint.** Energy of the photon = E = 1.02 MeV $= 1.02 \times (1.6 \times 10^{-13})\ J$.

$$\therefore \lambda = \frac{hc}{E} = 0.012\,\text{Å};\ d\lambda = 0.048\,\text{Å}$$

$\therefore$ Wavelength of scattered photon = λ' = 0.060 Å

Hence the energy of the scattered photon $= \frac{hc}{\lambda'} = 0.204$ MeV]

21. Both photoelectric and Compton effects arise due to the action of photons on electrons, but the two effects are not same. Explain this.

[**Ans.** In the photoelectric effect, the photon completely disappears and all of its energy is given to the photoelectron. In Compton effect, only a part of the energy of the incident photon is given to the Compton electron. The energy of the scattered photon is less than the energy of the incident photon.]

22. Explain why Compton effect is experimentally not observed for visible light rays.

[**Ans.** $\Delta\lambda = \frac{h}{m_0 c}(1-\cos\theta) = 0.024\,(1-\cos\theta)$. $\Delta\lambda$ depends only on the scattering angle θ. Maximum value of $\cos\theta = -1$ when θ = 180°. Hence the maximum value of $\Delta\lambda$ = 0.048 Å. This means that Compton effect can be detected only for those radiations whose wavelength is not greater than a few Å. For visible light ($\lambda \approx 5000$ Å), $(\Delta\lambda)_{\max}$ is only about 0.001% of the initial wavelength which cannot be detected.]

23. A photon recoils back after striking an electron at rest. What is the change in the wavelength of the photon? [**Ans.** $\Delta\lambda$ = 0.048 Å]

24. Explain the presence of unmodified line in Compton scattering.

[**Hint.** When a photon collides with a bound electron, its wavelength does not change.]

25. Explain why the Compton shift of wavelength is independent of the scattering material.

26. An X-ray is found to have its wavelength doubled on being scattered through 90°. Calculate the wavelength with $c = 3 \times 10^8\ m/s$, $h = 6.62 \times 10^{-34}\ Js$ and $m_0 = 9.1 \times 10^{-31}\ kg$. [**Ans.** 0.0242 Å]

27. X-rays of wavelength 1.0×10^{-11} m are incident on free stationary electrons. What is the wavelength of the X-rays that are scattered directly backwards ? What is the energy gained by the scattered electrons in this case ?

[**Hint.** $d\lambda = 0.48 \times 10^{-11}$ m. $\lambda' = \lambda + d\lambda = 1.48 \times 10^{-11}\ m$. The incident and scattered photons have energies $h\nu$ and $h\nu'$. The scattered electron must therefore receive the difference in energy: $U = h\nu - h\nu' = h\left(\frac{c}{\lambda} - \frac{c}{\lambda'}\right) = 40\ keV$].

28. Find the Miller indices of a set of parallel planes which make intercepts in the ratio $3a : 4b$ on the X and Y axes and are parallel to the Z axis a, b, c being primitive vectors of the lattice. [**Ans.** (430)]

29. In an orthorhombic crystal, a lattice plane cuts intercepts of lengths $3a$, $-2b$ and $3c/2$ along three axes. Deduce the Miller indices of the plane. a, b, c are primitive vectors of the unit cell.

[**Hint.** $ra : sb : tc = 3a : (-2b) : 3c/2$ or $r : s : t = 3 : (-2) : 3/2$

or $\frac{1}{r} : \frac{1}{s} : \frac{1}{t} : = \frac{1}{3} : \left(-\frac{1}{2}\right) : \frac{2}{3}$.

Multiplying by 6, $\frac{1}{r} : \frac{1}{s} : \frac{1}{t} = 2 : (-3) : 4$.

Hence Miller indices of the plane are $(2, \bar{3}, 4)$.]

30. In a simple cubic crystal, find the ratio of intercepts on the three axes by the (123) plane. [**Ans.** 6 : 3 : 2]

31. Calculate the interplanar spacing for a (321) plane in a simple cubic lattice whose lattice constant is 4.2×10^{-10} m.

[**Hint.** $d_{hkl} = \dfrac{a}{\sqrt{h^2 + k^2 + l^2}}$

$$\therefore \quad d_{321} = \frac{4.2 \times 10^{-10}}{\sqrt{3^2 + 2^2 + 1^2}} = 1.1 \times 10^{-10} \ m]$$

32. X-rays of wavelength 0.71 Å are reflected from the (110) plane of a rocksalt crystal (a = 2.82 Å). Calculate the glancing angle corresponding to second order reflection.

[**Hint.** $d_{110} = \dfrac{a}{\sqrt{2}} = \dfrac{2.82}{\sqrt{2}} = 2\text{Å}$. Now, $2d \sin\theta = n\lambda$

$d = 2\text{Å}$, $n = 2$ and $\lambda = 0.71$ Å.

$\therefore \quad \theta = 20°42'$

33. Explain why the X-ray spectra of elements of nearby atomic numbers are qualitatively very similar, while the optical spectra of these elements may differ considerably.

[**Ans.** The transitions that give rise to X-ray spectra are the same in all elements, since these transitions involve only inner, closed-shell electrons. The transitions that give rise to optical spectra involve the outermost electrons. Hence the possible transitions depend critically upon the degree of occupancy of the outermost shell.]

34. An X-ray beam of energy 0.01 MeV is reflected at the (100) plane of sylvine crystal (d_{100} = 0.314 nm). Calculate the glancing angle θ at which the first order Bragg's spectrum will be observed.

[**Hint.** $\lambda = \dfrac{hc}{E} = \dfrac{6.625 \times 10^{-34} \times 3 \times 10^{8}}{0.01 \times 1.6 \times 10^{-13}} = 1.242 \times 10^{-10}$ m.

$2d \sin\theta = n\lambda \quad \therefore \ \theta = 10° \ 31']$

35. Define "Bravais lattice".

[**Ans.** We give two equivalent definitions of a Bravais lattice:

(*a*) A Bravais lattice is an infinite array of discrete points with an arrangement and orientation that appears *exactly* the same, from whichever of the points the array is viewed.

(*b*) A (three-dimensional) Bravais lattice consists of all points with position vectors **R** of the form

$$\mathbf{R} = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3, \qquad \text{...(1)}$$

Here, $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ are any three vectors not all in the same plane. n_1, n_2, and n_3 range through all integral values.

Thus the point $\Sigma n_i \mathbf{a}_i$ is reached by moving n_i steps of length $\mathbf{a}_i$ in the direction of $\mathbf{a}_i$ for i = 1, 2 and 3.

The vectors $\mathbf{a}_i$ are called *primitive vectors* and are said to *generate* or *span* the lattice.]

Example. Fig. 7.35 shows a simple cubic three-dimensional Bravais lattice. The three primitive vectors can be taken to be mutually perpendicular, and with a common magnitude.

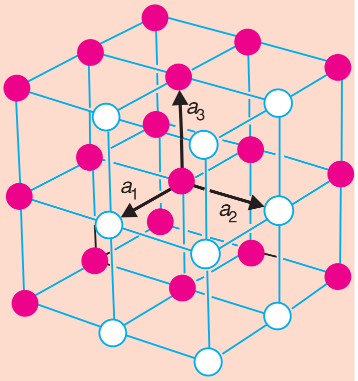


Fig. 7.35

CHAPTER 8

THE PHOTOELECTRIC EFFECT

AT A GLANCE

8.1 Introduction
8.2 Lenard’s Method to Determine e/m for Photoelectrons
8.3 Richardson and Compton Experiment
8.4 Experimental Investigations on the Photoelectric Effect
8.5 Einstein’s Photoelectric Equation
8.6 Photoelectric Cells

8.1 Introduction

Whenever light or electromagnetic radiations (such as X-rays, Ultraviolet rays) fall on a metal surface, it emits electrons. *This process of emission of electrons from a metal plate, when illuminated by light of suitable wavelength, is called the photoelectric effect.* The electrons emitted are known as the photoelectrons. In the case of alkali metals, photoelectric emission occurs even under the action of visible light. Zinc, cadmium etc., are sensitive to only ultraviolet light.

The Nature of Photo-particles

The arrangement used by Hallwachs is shown in Fig. 8.1. The apparatus consists of two plates *A* and *C* placed in an evacuated quartz bulb.

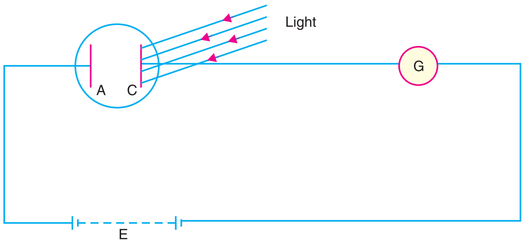


Fig. 8.1

The galvanometer (*G*) and battery (*E*) are connected as shown. When ultraviolet light is incident on the negative plate *C*, a current flows in the circuit as indicated by the galvanometer. But when light falls on the positive plate *A*, there is no current in the circuit. These observations show that photo-particle must be negatively charged.

8.2 Lenard's method to determine e/m for Photoelectrons

The apparatus used is shown in Fig. 8.2. It consists of a glass tube *G* which can be evacuated through the side tube *T*. Ultraviolet light passes through a quartz window *W* and falls on an aluminium cathode *C* enclosed in *G*. An earthed metal screen *A* with a small central hole forms the anode. The cathode (*C*) can be maintained at a desired potential, positive or negative relative to the anode *A*. P_1 and P_2 are small metal electrodes connected to electrometers E_1 and E_2.

When *C* is raised to a negative potential and illuminated, negatively charged particles are produced and accelerated towards the anode. A few particles pass through the hole in *A* and proceed with uniform velocity to P_1. Their arrival at P_1 is indicated by E_1. By applying a uniform magnetic field *B* (represented by the dotted circle) perpendicular to the plane of the figure and directed towards the reader, the photoelectrons can be deflected towards P_2. Their arrival at P_2 is indicated by the deflection they produce in E_2.

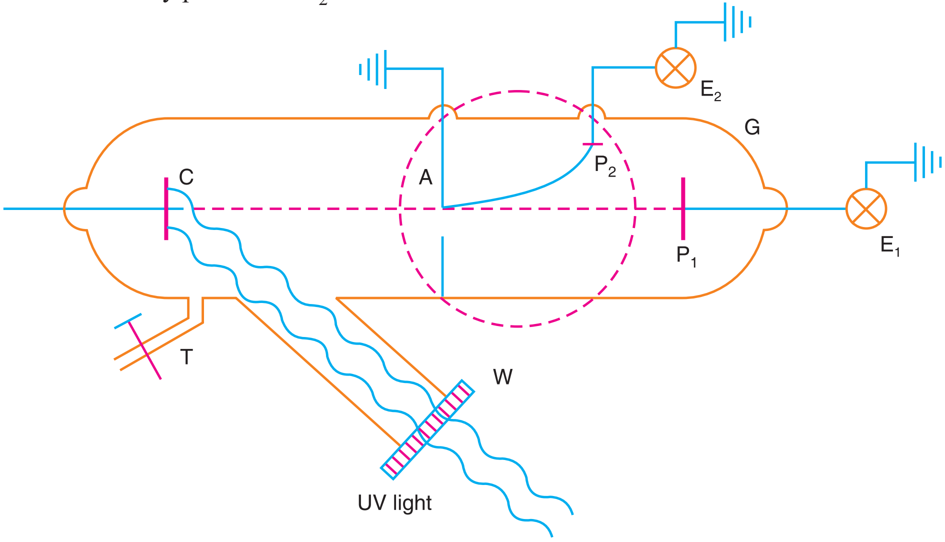


Fig. 8.2

Lenard first studied the relation between current and the potential applied to *C*. When the cathode potential was several volts positive, the current was zero. When *V* was +2 volts, there was a feeble current showing that a few particles possessed enough velocity to overcome the retarding potential of 2 volts. When the potential was further decreased, the current increased and reached a saturation value for –20 volts. Fig. 8.3 shows the variation of photoelectric current with cathode potential.

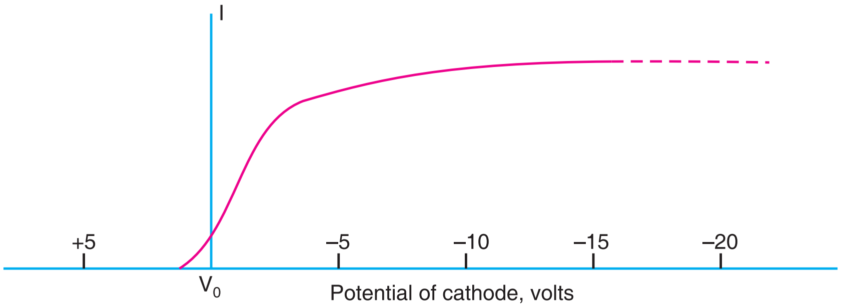


Fig. 8.3

After this preliminary investigation, Lenard applied to *C* a negative potential *V*, very large compared to the potential of 2 volts. The velocity imparted by the accelerating potential is so large that the velocity of the particles in the act of emission is negligible in comparison to it. Let *V* be the accelerating potential and *v* the velocity acquired by the photoelectrons. Then,

$$\frac{1}{2}mv^2 = eV \qquad \text{...(1)}$$

where *e* is the charge and *m* the mass of the photoelectron.

Let *R* be the radius of the circular path described by the photoelectrons in the region of uniform magnetic field of strength *B*.

$$\text{Then} \qquad \frac{mv^2}{R} = Bev.$$

$$\therefore \qquad v = \frac{BeR}{m} \qquad \text{..(2)}$$

Substituting this value of *v* in eqn. (1), $\frac{1}{2}m\,(BeR/m)^2 = eV$

$$\therefore \qquad \frac{e}{m} = \frac{2V}{B^2R^2} \qquad \text{...(3)}$$

Knowing *V*, *B* and *R*, *e*/*m* is calculated. Lenard found the value of *e*/*m* to be the same as that for electrons. *This clearly shows that the photoparticles are nothing but electrons.*

8.3 Richardson and Compton Experiment

The apparatus used by them is shown in Fig. 8.4. The emitter of photoelectrons (*C*) is a small strip of the metal under study and is kept at the centre of a spherical glass bulb *B*. *B* is silvered on the inner side and can be evacuated through the tube *T*. The silver coating on the inside of the bulb serves as the anode *A* and is connected to an electrometer *E* which measures the photoelectric current. Monochromatic light *L* is made to pass through a quartz window *W* and fall on *C*. *C* can be maintained at any desired potential *V* relative to *A* and this potential can be read with a voltmeter.

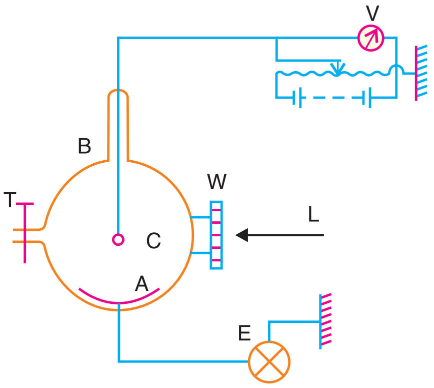


Fig. 8.4

(*i*) Relation between photoelectric current and retarding potential. They first studied the ralation between the photoelectric current *I* and the retarding potential *V*. Irradiating the cathode *C* with monochromatic light of a given intensity, the cathode potential (*V*) was varied from a few volts positive upto zero and negative values. The photoelectric current *I* was measured for different values of *V*. The observations were repeated with the intensity of illumination (*X*) doubled, trebled, etc. The relation between *I* and *V* is shown in Fig. 8.5. For a given intensity of illumination (X_1), there is no photoelectric current when the positive potential on the cathode is greater than a critical value + V_0. At potentials just less than + V_0, a small current is produced. As the potential decreases to zero, the current rapidly increases and reaches a maximum value when $V = 0$. No further increases

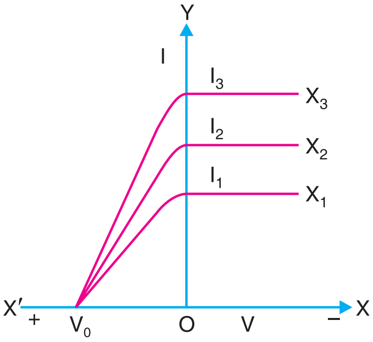


Fig. 8.5

in current are observed when V becomes negative. Hence I_1 is the maximum current due to illumination of intensity X_1. When the intensity is doubled (X_2) as in (2), the maximum current I_2 is double I_1: But the critical potential V_0 is the same as before. When the intensity of illumination is increased three fold (X_3), the corresponding maximum current increases proportionately, but the critical potential remains unchanged. Thus, the maximum current I_m is proportional to the intensity of illumination X, *i.e.*, $l_m \propto X$. The critical potential V_0 is independent of the intensity.

(*ii*) Relation between velocity of photoelectrons and the frequency of light. Several monochromatic radiations of wavelengths λ_1, λ_2, λ_3, etc., are allowed to fall on the emitter. The intensity of illumination for each wavelength is adjusted to give the same value of I_m in each case. In each case, the photoelectric current I for different values of V is determined. Fig. 8.6 represents the nature of the curves. It is seen from the curves that if $\lambda_1 > \lambda_2 > \lambda_3$ the critical potentials are $V_1 < V_2 < V_3$. Hence, as the wavelength of light increases, the critical retarding potential decreases. This means that the maximum K.E. of the photoelectrons (given by $\frac{1}{2} mv_m^2 = eV_0$) increases with increasing frequency of light. Since V_0 is independent of the intensity of illumination, we conclude that the velocity and K.E. of photoelectrons are independent of the intensity of illumination, but dependent on the frequency of the incident light.

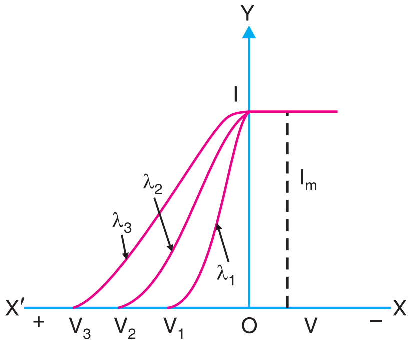


Fig. 8.6

A linear relation is found to exist between the maximum energy of emission (eV_0) and the frequency (ν) of the light. If eV_0 is plotted against ν, we get, for any emitting surface a straight line (Fig. 8.7), whose intercept on the frequency axis gives the threshold frequency $\nu_0 = c/\lambda_0$ for this surface. The slope of the straight line is the same for all emitting surfaces and is found to be equal to h. The value of ν'_0 however, depends on the emiting surface. ν_0 is called the threshold frequency, as it represents the beginning of the photoelectric activity of the emitter. *Threshold frequency is defined as the minimum value of frequency of incident light below which the photo-electric emission stops completely, howsoever high the intensity of light may be.* At threshold frequency, the kinetic energy of emitted photo-electrons is just zero.

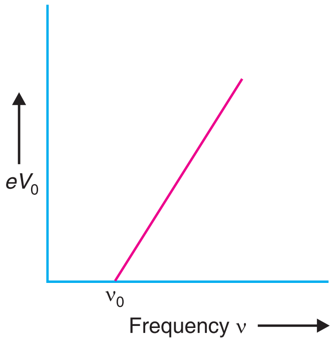


Fig. 8.7

8.4 Experimental Investigations on the Photoelectric Effect

Photoelectric effect can be studied in detail with the help of the apparatus shown in Fig. 8.8.

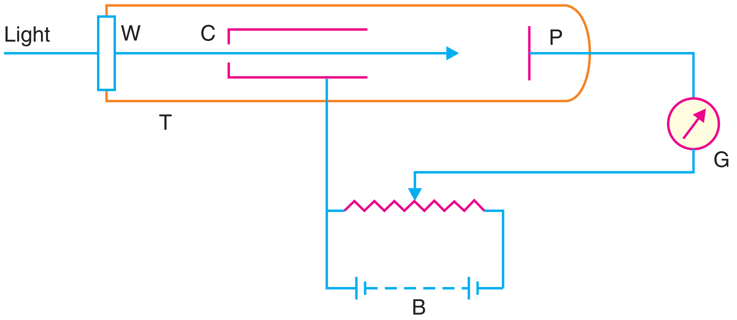


Fig. 8.8

Photoelectric Effect with Projected Electroscope.

It consists of an evacuated glass tube *T* with a quartz window *W*. *P* is a photoelectrically sensitive plate. *C* is a hollow cylinder and it has a small hole that permits the incident light to fall on the plate *P*. *P* is connected to the negative end. *C* is connected to the positive terminal of a battery *B*. When light from some source falls on the plate *P*, the photoelectrons are ejected out of the plate *P*. These photoelectrons are attracted by the positively charged cylinder *C*. Hence a photoelectric current flows from *P* to *C* in the bulb and from *C* to *P* outside the bulb. This current can be measured from the deflection produced in the galvanometer *G*. It is found that the strength of the photoelectric current increases as the potential of *C* is more and more positive with respect to *P*. The deflection in *G* decreases when the potential of *C* is negative with respect to *P*. The results obtained can be summarised into four statements, which are known as the laws of photoelectric emission.

Laws of photoelectric emission. (*i*) *For every metal, there is a particular minimum frequency of the incident light, below which there is no photoelectric emission, whatever be the intensity of the radiation. This minimum frequency, which can cause photoelectric emission, is called the threshold frequency.*

(*ii*) *The strength of the photoelectric current is directly proportional to the intensity of the incident light, provided the frequency is greater than the threshold frequency.*

(*iii*) *The velocity and hence the energy of the emitted photoelectrons is independent of the intensity of light and depends only on the frequency of the incident light and the nature of the metal.*

(*iv*) *Photoelectric emission is an instantaneous process. The time lag, if any, between incidence of radiation and emission of the electrons, is never more than* 3×10^{-9} *sec.*

Failure of the electromagnetic theory. The above experimental facts could not be explained on the basis of the electromagnetic theory of light.

(1) Calculations showed that it would require about 500 days to dislodge a photoelectron from sodium by exposure to violet light of wavelength 4000 Å. Experimentally, however, we observe that electron ejection commences without delay.

(2) According to the classical theory, light of greater intensity should impart greater K.E. to the liberated electrons. But, this does not happen. Also, the velocity of the emitted electron should not depend on the frequency of the incident light. But it does.

The phenomenon was adequately explained by Einstein on the basis of Planck's Quantum theory of radiation.

Quantum theory. According to Planck, the energy of a monochromatic wave with frequency ν can only assume those values which are integral multiples of energy $h\nu$. *i.e.*, $E_n = nh\nu$, where n is an integer referring to the number of "*Photons*". Thus the energy of a single *PHOTON* of frequency ν is $E = h\nu$.

8.5 Einstein's Photoelectric Equation

According to Einstein, light of frequency ν consists of a shower of corpuscles or photons each of energy $h\nu$. When a photon of light of frequency ν is incident on a metal, the energy is completely transferred to a free electron in the metal. A part of the energy acquired by the electron is used to pull out the electron from the surface of the metal and the rest of it is utilised in imparting K.E. to the emitted electron. Let ϕ be the energy spent in extracting the electron from the emitter to which it is bound (photoelectric work function) and ½ mv^2 the K.E. of the photoelectron.

Then
$$h\nu = \phi + \frac{1}{2}mv^2 \quad \text{...(1)}$$

This relation is known as the *Einstein's Photoelectric equation.* If ν_o is the threshold frequency which just ejects an electron from the metal without any velocity then, $\phi = h\nu_o$.

$$\therefore \quad h\nu = h\nu_o + \frac{1}{2}mv_{\max}^2 \quad \text{...(2)}$$

where v_{max} is the maximum velocity acquired by the electron.

or
$$\frac{1}{2}mv^2_{max} = h(\nu - \nu_0) \quad \text{...(3)}$$

Notes

(*i*) *The work function of a metal may be defined as the energy which is just sufficient to liberate electrons from the metal surface with zero velocity.*

(*ii*) Equation (3) suggests that the energy of the emitted photoelectrons is independent of the intensity of the incident radiation but increases with the frequency.

Experimental verification of Einstein's Photoelectric Equation—Millikan's Experiment.

Theory. Millikan's experiment is based on what is known as the "*stopping potential*". The stopping potential is the necessary retarding potential difference required in order to just halt the most energetic photoelectron emitted.

$$\left.\begin{array}{l}\text{The K.E. of a photoelectron leaving}\\ \text{the surface of a metal irradiated with}\\ \text{light of frequency } \nu\end{array}\right\} = \frac{1}{2}mv_{\max}^2 = h\nu - \phi.$$

Let V be the P.D. which is applied between the emitter and a collecting electrode in order to prevent the photoelectron from just leaving the emitter, the emitter being maintained at a positive potential with respect to the collector. Then,

$$eV = \frac{1}{2}mv_{\max}^2$$

$$\therefore \quad eV = h\nu - \phi$$

or $$V = \frac{h}{e}\nu - \frac{\phi}{e} \qquad \text{...(1)}$$

ϕ is constant for a given metal; h and e are also constants.

Hence, Equation (1) represents a straight line. V is measured for different values of ν. A graph is then plotted between the stopping potential (V) taken along the Y-axis and the frequency of light (ν) taken along the X-axis. The graph is a straight line (Fig. 8.9). The slope of the straight line

$$\tan\theta = \frac{h}{e}$$

$$\therefore \qquad h = e\tan\theta \qquad \text{...(2)}$$

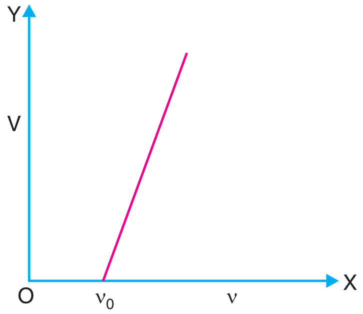


Fig. 8.9

Hence the value of h (Planck's constant) can be calculated. The intercept on the X-axis gives the threshold frequency ν_o for the given emitter. From this, photoelectric work function = $\phi = h\ \nu_o$ can be calculated.

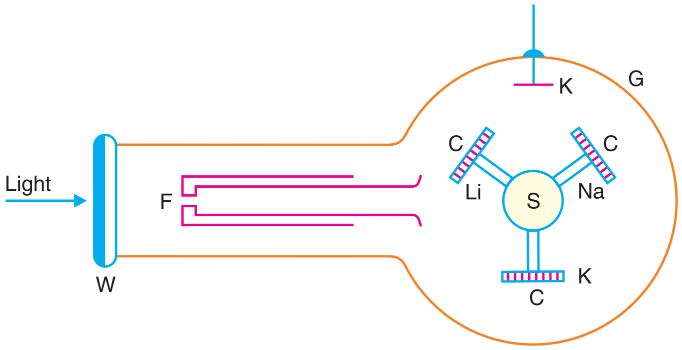


Fig. 8.10

Experiment. Millikan's apparatus is shown in Fig. 8.10. Alkali metals are employed as emitters, since they readily exhibit photoelectric emission even with visible light. Cylindrical blocks (C) of sodium, potassium or lithium are mounted on a spindle S at the centre of the glass flask G. The flask is evacuated to a very high vacuum to free the metals from all absorbed gases and to prevent their oxidation. The spindle can be rotated from outside by an electromagnet. As each metal block passes by the adjustable sharp edge K, a thin layer of it is removed, thus exposing a fresh surface of the metal to the irradiating light entering the flask through a quartz window W. Monochromatic light provided by a spectroscope is used to illuminate the fresh metal surfaces. The photoelectrons are collected by a Faraday cylinder F. The Faraday cylinder is made of copper oxide which is not photosensitive. The photocurrent is measured by an electrometer connected to the Faraday cylinder.

Robert Andrews Millikan (1868–1953)

As a Scientist, Millikan made numerous discoveries, chiefly in the fields of electricity, optics, and molecular physics. His earliest major success was the accurate determination of the charge carried by an electron, using the elegant "falling-drop method"; he also proved that this quantity was a constant for all electrons (1910), thus demonstrating the atomic structure of electricity. Next, he verified experimentally Einstein's all-important photoelectric equation, and made the first direct photoelectric determination of Plancks' constant h (1912-1915).

The stopping potential of the liberated photoelectrons is measured by raising the emitter surface to a positive potential, just sufficient to prevent any of the electrons from reaching the collector (F). The stopping potential is the positive potential applied to the emitter, which corresponds to zero current in the electrometer. The stopping potential (V) is determined for different wavelengths of the incident light. The value of V should be corrected for any contact potential between the metal (C) and Faraday cylinder (F). On plotting V against ν, we get a straight line. Measuring the slope of the straight line, the value of h/e is obtained. Then substituting the known value of e, h is calculated. The value of h calculated in this way agrees fairly well with the value obtained by other methods. Thus the Einstein's equation can be verified experimentally.

Example 1. *Calculate the work function of sodium, in electron-volts, given that the threshold wavelength is* 6800 Å, *and* $h = 6.625 \times 10^{-34}$ *Js.*

Sol. $\phi = h\nu_0 = hc/\lambda_0.$

Here, $h = 6.625 \times 10^{-34}$ Js; $c = 3 \times 10^8$ ms^{-1} and

$$\lambda_0 = 6800 \times 10^{-10} \text{ m.}$$

$$\therefore \quad \phi = \frac{(6.625 \times 10^{-34})\,(3 \times 10^8)}{(6800 \times 10^{-10})} \text{ J} = \frac{(6.625 \times 10^{-34})\,(3 \times 10^8)}{(6800 \times 10^{-10})\,(1.6 \times 10^{-19})}\, eV$$

$$= 1.827\ eV.$$

Example 2. *The photoelectric threshold for a metal is* 3000 Å. *Find the kinetic energy of an electron ejected from it by radiation of wavelength* 1200 Å.

Sol. Here, $\lambda_0 = 3000 \times 10^{-10}$ m; $\lambda = 1200 \times 10^{-10}$ m; $h = 6.62 \times 10^{-34}$ Js and $c = 3 \times 10^8$ ms^{-1}.

$$\text{K.E. of the electron} = \ = \frac{1}{2} mv^2 = h(\nu - \nu_0)$$

$$= h\left(\frac{c}{\lambda} - \frac{c}{\lambda_0}\right) = hc\left(\frac{1}{\lambda} - \frac{1}{\lambda_0}\right) = \frac{hc\,(\lambda_0 - \lambda)}{\lambda\lambda_0}$$

$$= \frac{(6.62 \times 10^{-34})\,(3 \times 10^8)\,(1800 \times 10^{-10})}{(3000 \times 10^{-10})\,(1200 \times 10^{-10})}$$

$$= 9.93 \times 10^{-19} \text{ J} = 6.2\ eV.$$

8.6 Photoelectric Cells

Photoelectric cell is an arrangement to convert light energy into electrical energy. There are three types of photocells, photoemissive, photovoltaic and photoconductive.

(*i*) Photo-emissive Cell. This consists of a glass or quartz bulb (B) according as it is to be used with visible or ultraviolet light (Fig. 8.11). C is the silver cathode in the form of a semi-cylindrical plate. The anode (A) is a rod mounted vertically at the centre of the bulb and parallel to its axis. A positive potential of about 100 volts is applied to the anode, the negative being connected to the cathode through a galvanometer G. When light falls on the cathode C, electrons are ejected from the cathode. A small current

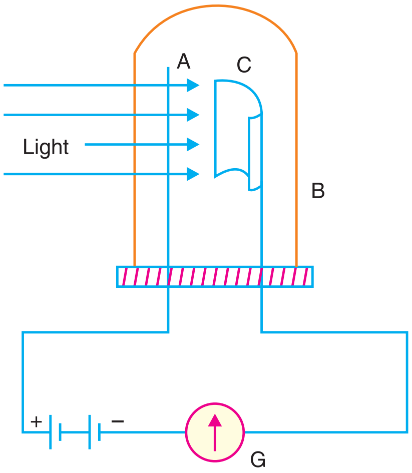


Fig. 8.11

flows through the cell and can be measured by the galvanometer. The photoemissive cell is used for reproduction of sound from photo-films.

Photoelectric Cells

(*ii*) Photo-voltaic cell. It consists of a layer of semiconductor material spread over a metallic surface by heat treatment. In one type of the photovoltaic cell, the metal plate is made of copper and the semiconductor is cuprous oxide (Cu_2O). On the other side of the semiconductor, there is a very thin layer of a translucent deposit which allows the semiconductor to be illuminated by radiations (Fig. 8.12). Light falling on the surface film (of gold or silver) penetrates into it and ejects photoelectrons from the semiconductor layer. These electrons travel in a direction opposite to the direction of the incident light. The conventional current, therefore, flows in the direction of the incident light. For small values of the resistance of the galvanometer, the current is directly proportional to the intensity of light. No external battery is required to operate a photovoltaic cell as the cell itself generates an e.m.f.

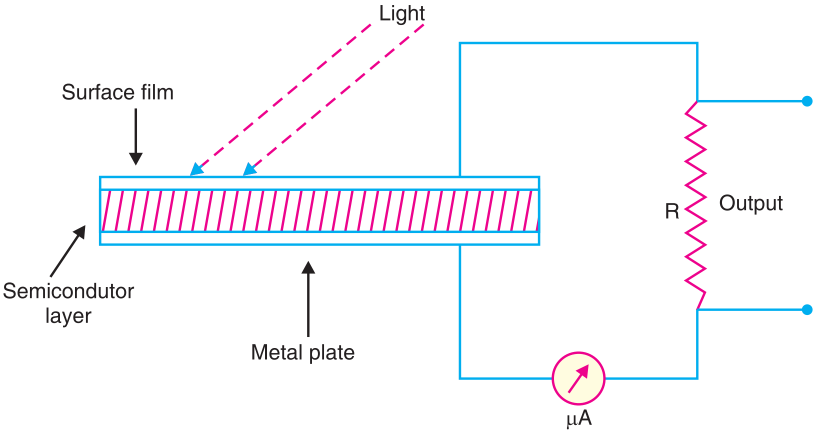


Fig. 8.12

(*iii*) Photoconductive Cell. These cells are based on the principle that the electrical resistance of a semiconductor material, like selenium, decreases with the increase of intensity of radiation incident upon it and conductivity is increased. A film of selenium is deposited on one side of an iron plate and a P.D. of 100 volts is applied between iron and selenium from an external battery. A galvanometer is included in the electric circuit. When a beam of light falls on the selenium film, a deflection will be observed in the galvanometer. As the intensity of the incident light is varied, the resistance of selenium also varies accordingly and the current in the circuit undergoes corresponding variations. The solar battery consists of several thousand photoconductive cells, which produce several kilowatt power.

Applications of Photoelectric cells. (*i*) *Exposure meters in photography.* An exposure meter is a device to calculate the correct time of exposure. The photoelectric cell in the instrument produces a current proportional to the light falling on it. The current operates a galvanometer, the scale of which is calibrated to read the time of exposure.

(*ii*) Photo-multiplier. It is based on the principle of *secondary emission.* When light strikes the surface of photosensitive metal plate *C*, it causes the ejection of photoelectrons from it (Fig. 8.13). These electrons are then attracted to a metal surface called a dynode, by setting a P.D. between the cathode *C* and the dynode 1. High energy electrons striking a metal surface can cause the ejection of one or more secondary electrons from the surface. Suppose that a photoelectron striking dynode 1 produces x electrons by secondary emission. These electrons are then directed towards dynode 2 by making its potential higher than that of dynode 1. Suppose x electrons are again ejected by secondary emission for each incident electron. Then, for each electron emitted by the photosensitive plate, there are now x^2 electrons and so on. If there are several dynodes, each at a potential higher than the preceding one, an avalanche of electrons reaches the collector plate A. A strong current then flows in the outer circuit. This device is used to amplify very weak light signals.

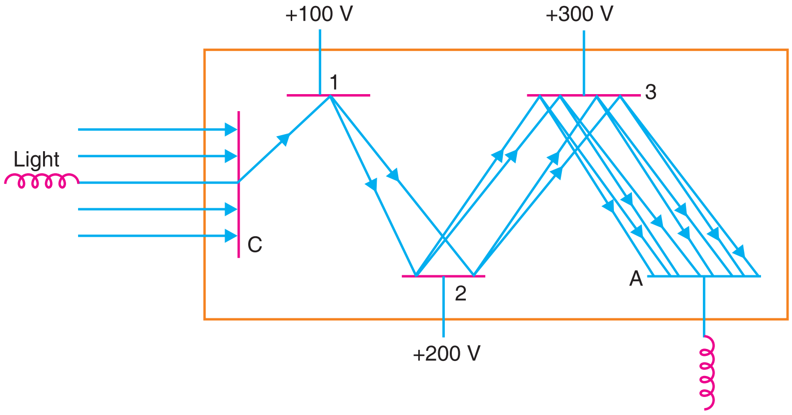


Fig. 8.13

(iii) Photoelectric cells are used to compare the illuminating powers of two light sources. They are also used in the measurement of the intensity of illumination of a light source.

***(iv)* Sound reproduction in films.** The film is provided with a sound track at one edge. Light passing through the sound track of the film falls on a photocell. Current is produced, which fluctuates correspondingly with the intensity of sound recorded in the sound track. The current impulses are converted to sound by speakers.

***(v)* Automatic operation of street lights.** A photoelectric cell, located in a street light circuit, switches off the street light when sunlight is incident on the cell. When sunlight fades and it becomes dark, the photoelectric cell switches on the street lights.

EXERCISE

1. What is photoelectric effect ? Describe simple experiments to study photoelectric phenomena. State the laws of photo-electric emission.

2. Explain photoelectric effect. Give an account of Einstein's explanation of photo-electric effect on the basis of quantum theory.

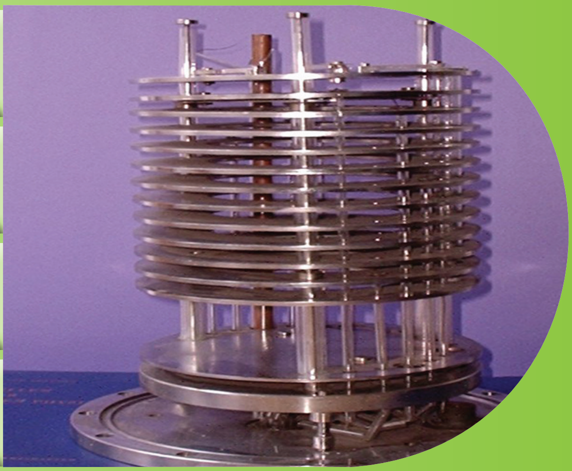
3. Derive Einstein's photo-electric equation and describe Millikan's experiment to verify the same.

4. Give an account of Millikan's experimental verification of Einstein's photoelectric equation. Explain how Millikan calculated the value of the Planck's constant and work function of the material.

5. Describe different types of photo-electric cells and explain their action. Mention some important uses of photo-electric cells.

6. Light of wavelength 4300 Å is incident on (*a*) nickel surface of work function 5 electron-volts and (*b*) a potassium surface of work function 2.3 electronvolts. Find out, if electrons will be emitted, and if so, the maximum velocity of the emitted electrons in each case.

[**Ans.** (*a*) $\lambda_0 = 2484$ Å; Electrons will not be emitted. (*b*) $\lambda_0 = 4389$ Å. As $\lambda_0 > \lambda$, the electrons will be emitted. $v = 1.423 \times 10^5$ ms^{-1}].



CHAPTER 9

PLANCK'S QUANTUM THEORY

The distribution of energy in the spectrum of a black body. If the radiation emitted by a black body at a fixed temperature is analysed by means of a suitable spectroscopic arrangement, it is found to spread up into a continuous spectrum. The total energy is not distributed uniformly over the entire range of the spectrum. The distribution of energy in various parts of the spectrum was experimentally studied by Lummer and Pringsheim. The radiation from the black body was rendered into a parallel beam by the concave mirror [Fig. 9.1]. It is then allowed to fall on a prism of fluorspar to resolve it into a spectrum. The spectrum is brought to focus by another concave mirror on to a linear bolometer. The bolometer is connected to a galvanometer. The deflections in the galvanometer corresponding to different λ are noted by rotating the prism table. Then curves are plotted for E_λ versus λ. The experiment is done with the black body at different temperatures. The curves obtained are shown in Fig. 9.2.

Results. (*i*) At any given temperature, E_λ first increases with λ, reaching a maximum value corresponding to a particular wavelength λ_m and then decreases for longer wavelengths.

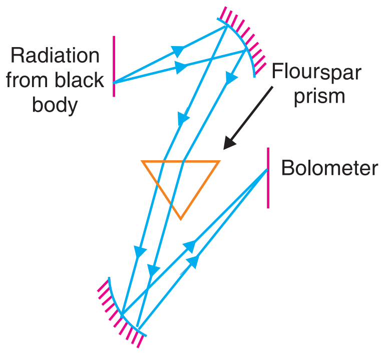


Fig. 9.1

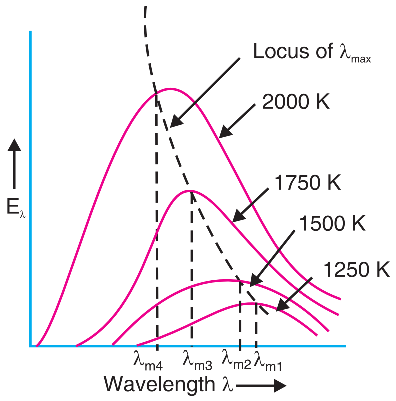


Fig. 9.2

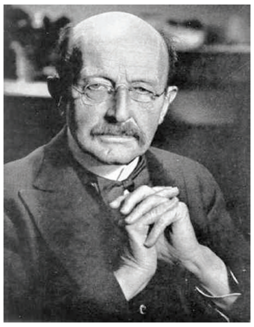
(*ii*) The value of E_λ for any λ increases as temperature increases.

(*iii*) The wavelength corresponding to the maximum energy shifts to shorter wavelength side as the temperature increases. This confirms Wien's displacement law $\lambda_m T = \text{constant}$.

(*iv*) Total energy emitted per unit area of the source per second at a given temperature is $\int_0^\infty E_\lambda \, d\lambda$. It will be represented by the total area between the curve for that temperature and the λ-axis. This area is found to be proportional to the fourth power of the absolute temperature. This verifies *Stefan's law*.

Wien's Displacement Law. *The wavelength of the most strongly emitted radiation in the continuous spectrum from a full radiator is inversely proportional to the absolute temperature of that body*, *i.e.*, $\lambda_m T = b$.

Here, b is Wien's constant 2.898×10^{-3} mK.



Max Planck (1858–1947)

Planck's hypothesis. According to the classical theory of radiation, energy changes of radiators take place continuously. The classical theory failed to explain the experimentally observed distribution of energy in the spectrum of a black body. Planck succeeded in deriving a formula which agrees extremely well with experimental results. He discarded both the idea of radiation being a continuous stream as well as the law of equipartition of energy. He suggested the *quantum theory of radiation*. His assumptions are:

(1) A black-body radiation chamber is filled up not only with radiation, but also with *simple harmonic oscillators* or *resonators* of molecular dimensions. They can vibrate with all possible frequencies.

(2) The oscillators or resonators cannot radiate or absorb energy continuously. But an oscillator of frequency ν can only radiate or absorb energy in units or quanta of magnitude $h\nu$. h is a universal constant called Planck's constant. The emission of radiation corresponds to a decrease and absorption to an increase in the energy and amplitude of an oscillator.

Derivation of Planck's law of radiation. Let N be the total number of Planck's resonators and E their total energy. Then, average energy per oscillator $= \bar{\varepsilon} = E/N$. Now,

$$N = N_0 + N_0 e^{-\varepsilon/kT} + N_0 e^{-2\varepsilon/kT} + \cdots + N_0 e^{-r\varepsilon/kT} + \cdots$$

Here N_0 = number of resonators having 0 energy.

$N_0 e^{-\varepsilon/kT}$ = number of resonators having energy ε,

$N_0 e^{-2\varepsilon/kT}$ = number of resonators having energy 2ε,

$N_0 e^{-r\varepsilon/kT}$ = number of resonators having energy $r\varepsilon$ and so on.

Putting $\varepsilon/kT = x$,

$$N = N_0 + N_0 e^{-x} + N_0 e^{-2x} + \cdots\cdots\cdots\cdots\cdots N_0 e^{-rx}$$

$$\therefore \quad N = \frac{N_0}{1-e^{-x}} \qquad \ldots(1)$$

The total energy of Planck's resonators is

$$E = 0 \times N_0 + \varepsilon \times N_0 e^{-x} + 2\varepsilon \times N_0 e^{-2x} + \cdots + r\varepsilon \times N_0 e^{-rx} + \cdots\cdot$$

$$Ee^{-x} = \varepsilon N_0 e^{-2x} + 2\varepsilon N_0 e^{-3x} + \cdots + r\varepsilon N_0 e^{-(r+1)x}$$

Subtracting, $E(1 - e^{-x}) = \varepsilon N_0 e^{-x} + \varepsilon N_0 e^{-2x} + \varepsilon N_0 e^{-3x} + \cdots$

$$= \frac{\varepsilon N_0 e^{-x}}{1-e^{-x}}$$

$$\therefore \qquad E = \frac{\varepsilon N_0 e^{-x}}{(1-e^{-x})^2} \qquad \text{... (2)}$$

$$\left.\begin{array}{l}\text{Average energy}\\ \text{of a resonator}\end{array}\right\} = \bar{\varepsilon} = \frac{E}{N} = \frac{\varepsilon e^{-x}}{1-e^{-x}} = \frac{\varepsilon}{e^x - 1}$$

According to Planck's hypothesis, $\varepsilon = h\nu$. Further $\nu = c/\lambda$.

Hence,

$$\varepsilon = \frac{hc}{\lambda} \quad \text{and} \quad x = \frac{\varepsilon}{kT} = \frac{hc}{\lambda kT}$$

$$\therefore \qquad \bar{\varepsilon} = \frac{hc/\lambda}{\left(e^{hc/\lambda kT} - 1\right)} \qquad \text{...(3)}$$

Number of oscillators per unit volume in the wavelength range λ and $\lambda + d\lambda = 8\pi\,\lambda^{-4}\, d\lambda$. ...(4)

Hence, energy density of radiation between wavelengths λ and $\lambda + d\lambda$ = (average energy of a Planck's oscillator) × (number of oscillators per unit volume). Therefore,

$$E_\lambda\, d\lambda = \frac{hc/\lambda}{\left(e^{hc/\lambda kT} - 1\right)} \times 8\pi\lambda^{-4}\, d\lambda = \frac{8\pi hc\lambda^{-5}}{\left(e^{hc/\lambda kT} - 1\right)} d\lambda \qquad \text{...(5)}$$

$$\text{or} \qquad E_\nu d\nu = \frac{\pi \;\; \nu}{c\,\left(e^{h\nu/kT} - 1\right)} \qquad \text{...(6)}$$

Here $E_\nu\, d\nu$ is the energy density belonging to the range $d\nu$.

Eq. (5) represents Planck's radiation law in terms of wavelength. Planck's formula fits the experimental curve very closely [Fig. 9.3].

Planck's formula reduces to Wien's formula for small wavelengths. When λ is small, $e^{hc/\lambda kT}$ is large when compared to 1. Hence Eq. (5) reduces to

$$E_\lambda d\lambda = 8\pi\, h\, c\, \lambda^{-5}\, e^{-hc/\lambda kT}\, d\lambda \qquad \text{...(7)}$$

This is *Wien's law*.

Planck's formula reduces to Rayleigh Jean's formula for longer wavelengths.

When λ is large, $e^{-hc/\lambda kT} \approx 1 + (hc/\lambda kT)$.

Hence Planck's law reduces to

$$E_\lambda\, d\lambda = 8\pi \frac{hc\lambda^{-5}}{(hc/\lambda kT)} d\lambda = 8\pi kT\lambda^{-4}\, d\lambda \qquad \text{...(8)}$$

This is *Rayleigh-Jeans formula*.

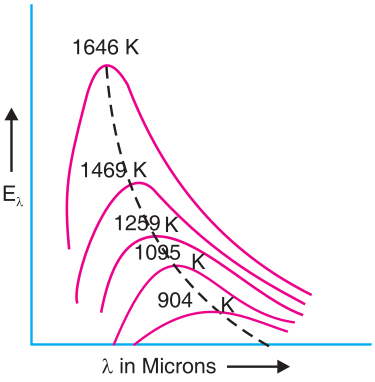


Fig. 9.3

CHAPTER 10

ELECTRON

AT A GLANCE

10.1 Dunnington's Method for Determining e/m

10.1 Dunnington's Method for Determining e/m

The apparatus is shown in Fig. 10.1. An alternating voltage at a constant high frequency produced by a crystal oscillator O is applied simultaneously to the two pairs of electrodes AA_1 and BB_1. Electrons from the hot filament F are accelerated towards B_1 during the positive half-cycle and emerge through a fine opening in B_1. The electrons are then bent into a circular path by a magnetic field B applied normal to the plane of the figure. The radius of curvature r of the circular path is defined by the slits S_1, S_2 and S_3. Then only those electrons whose velocity satisfies the relation $Bev = mv^2 / r$ can pass through the slits S_1, S_2 and S_3.

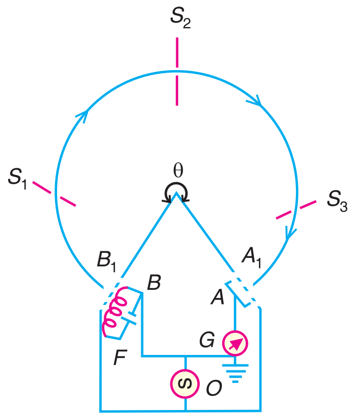


Fig. 10.1

∴ the speed of the electron in its circular path

$$v = Ber / m \qquad \text{...(1)}$$

After the electrons are turned through an angle θ, they enter the Faraday chamber A through a grid A_1 and produce a deflection in the galvanometer G. The electrodes and the slits are enclosed in an evacuated glass envelope. The grid A_1 is connected to the same oscillator which accelerates the electrons. Suppose the time taken by the electrons to be turned through the angle θ is the period $1/f$ of the oscillator or an integral multiple of it, say n/f. Then the electrons will lose all their energy in overcoming the opposing P.D. between A_1 and A and will just fail to reach A. The galvanometer indicates zero deflection when this condition is satisfied. This can be brought about by varying the intensity of the magnetic field B.

Now, distance travelled by the electron in going from B_1 to $A_1 = r\,\theta$.

The time taken for travelling this distance $= n/f$.

$$\therefore \qquad v = \frac{\text{Distance}}{\text{Time}} = \frac{r\theta}{n/f} = \frac{r\theta f}{n} \qquad \text{...(2)}$$

Hence, from (1) and (2),

$$\frac{Ber}{m} = \frac{r\theta f}{n} \text{ or } \frac{e}{m} = \frac{f\theta}{nB}.$$

θ is measured using microscopes built into the apparatus. The magnetic field (B) can be determined accurately. f can be measured with an accuracy of one part in one million. This method is, therefore, a precision one. Thus, e/m is calculated. The value obtained was $e/m = 1.7597 \times 10^{11}$ coulomb per kg.

Example 1. *In a determination of e/m by Dunnington's method, frequency of A. C. = 20 MHz and θ = 315°. If the successive values of B for which the current is minimum are* 1.617×10^{-4} *and* 1.293×10^{-4} *weber /* m^2*, find the value of e/m.*

Sol. We know that, $\dfrac{e}{m} = \dfrac{f\theta}{nB}$ where n is an integer.

Here,
$$f = 20 \times 10^6 \text{ Hz};$$
$$\theta = 315° = 315 \times \frac{\pi}{180} \text{ radians} = \frac{7\pi}{4} \text{ radians}.$$

Now
$$\frac{e}{m} = \frac{(20 \times 10^6)(7\pi/4)}{n \times (1.617 \times 10^{-4})} = \frac{(20 \times 10^6)(7\pi/4)}{(n+1)(1.293 \times 10^{-4})}$$

$$\therefore \qquad n \times 1.617 \times 10^{-4} = (n+1)\, 1.293 \times 10^{-4}$$

or $$1.617n = 1.293n + 1.293 \quad \text{or} \quad 0.324n = 1.293$$

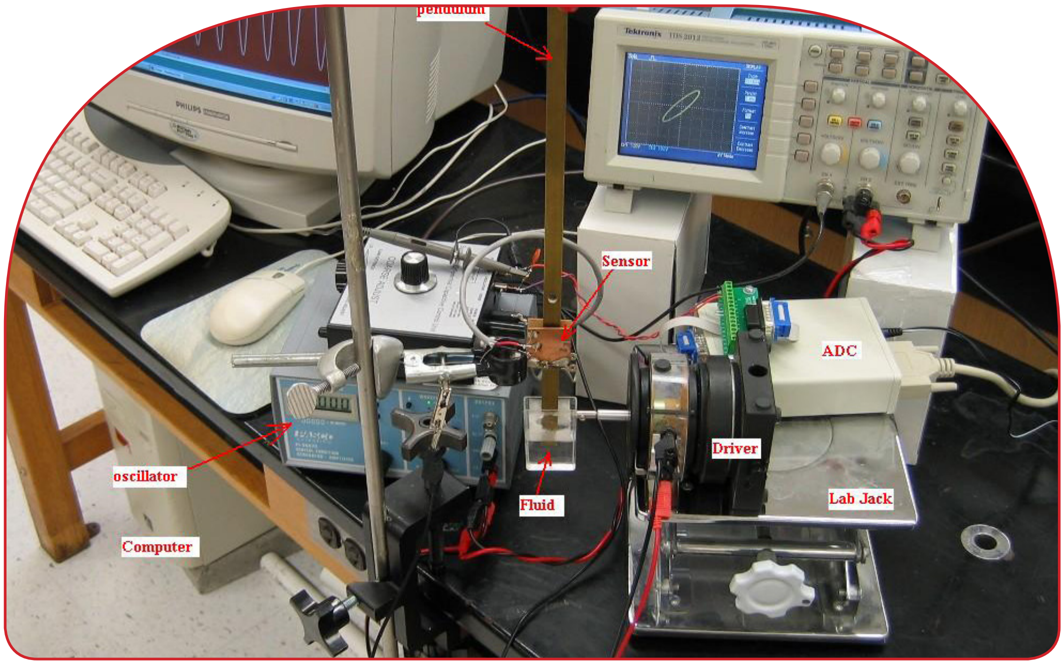
or $$n = 1.293 / 0.324 = 4 \text{ (We take the nearest integer).}$$

$$\therefore \qquad \frac{e}{m} = \frac{(20 \times 10^6)(7\pi/4)}{4 \times (1.617 \times 10^{-4})} = 1.7 \times 10^{11}\,\text{C kg}^{-1}.$$

PART – III

QUANTUM MECHANICS

11. Wave Mechanics
12. Quantum Mechanics II
13. Quantum Mechanics III
14. Quantum Mechanics IV
15. Quantum Mechanics V
16. Quantum Mechanics VI
17. Three Dimensional Harmonic Oscillator
18. Identical Particles and Spin


pendulum
Sensor
ADC
Driver
oscillator
Fluid
Lab Jack
Computer

CHAPTER 11

Wave Mechanics

AT A GLANCE

11.1 Introduction
11.2 Expression for Group Velocity
11.3 Experimental Study of Matter Waves
11.4 Heisenberg's Uncertainty Principle
11.5 Wave Mechanical Atom Model
11.6 Mathematical Proof of Uncertainty Principle for One Dimensional Wave-packet
11.7 Basic Postulates of Wave Mechanics
11.8 Derivation of Time-dependent form of Schrodinger Equation
11.9 Properties of the Wave Function
11.10 The Particle in a Box – Infinite Square Well Potential
11.11 Potential Step
11.12 The Barrier Penetration Problem
11.13 Linear Harmonic Oscillator
11.14 The Hydrogen Atom
11.15 The Rigid Rotator

11.1 Introduction

According to de Broglie a *moving particle, whatever its nature, has wave properties associated with it.* He proposed that the wavelength λ associated with any moving particle of momentum p (mass m and velocity v) is given by

$$\lambda = \frac{h}{p} = \frac{h}{mv}, \qquad ...(1)$$

where h is Planck's constant. Such waves associated with the matter particles are called *matter waves* or *de Broglie waves.*

Bohr's Theory of the hydrogen atom led de Broglie to the conception of matter waves. According to Bohr's theory, the stable states of electrons in the atom are governed by "*integer rules*". The only phenomena involving integers in physics are those of interference and modes of vibration of stretched strings, both of which imply wave motion. Hence de Broglie thought that the electrons may also be characterised by a periodicity. So he proposed that *matter*, *like radiation*, *has dual nature.* Eq. (1) was verified by experiments involving the diffraction of electrons by crystals.

Louis de Broglie (1892-1987)

The de Broglie Wavelength

A photon of light of frequency ν has the momentum

$$p = h\nu/c.$$

But $\nu = c/\lambda$. Therefore, the momentum of the photon can be expressed in terms of wavelength λ as

$$p = h/\lambda.$$

The wavelength of a photon is, therefore, specified by its momentum according to the relation

$$\lambda = h/p. \qquad \text{...(1)}$$

de Broglie suggested that Eq. (1) is a completely general one that applies to material particles as well as to photons. The momentum of a particle of mass m and velocity v is $p = mv$, and its de Broglie wavelength is accordingly

$$\lambda = h/mv.$$

If E_k is the kinetic energy of the material particle, then

$$p = \sqrt{(2mE_k)}$$

Therefore, de Broglie wavelength of particle of K.E. = E_k is given by

$$\lambda = \frac{h}{\sqrt{2mE_k}} \qquad \text{...(2)}$$

If a charged particle carrying charge q is accelerated through a potential difference V volts, then kinetic energy $E_k = qV$.

∴ *The de Broglie wavelength for charged particle* of charge q and accelerated through a potential difference of V volts is

$$\lambda = \frac{h}{\sqrt{2mqV}} \qquad \text{...(3)}$$

EXAMPLE 1. *Find the de Broglie wavelength associated with*

(*i*) *A 46 gm golf ball with velocity 36 m/s.*

(*ii*) *an electron with a velocity* 10^7 *m/s.*

Which of these two show wave character and why ? **(Garhwal 1994)**

SOL. (*i*) Since $v << c$, we can take $m = m_0 \rightarrow$ the rest mass. Hence

$$\lambda = \frac{h}{mv} = \frac{6.63 \times 10^{-34}\text{ J.s}}{(0.046\text{ kg})(36\text{ m/s})}$$

$$= \mathbf{4.0 \times 10^{-34}\ m.}$$

Thus the wavelength associated with golf ball is much smaller as compared to its dimensions. Hence **no wave aspects** can be expected in its behaviour.

(*ii*) Again $v << c$, so $m = m_0 = 9.1 \times 10^{-31}$ kg.

$$\therefore \qquad \lambda = \frac{h}{mv} = \frac{6.63 \times 10^{-34}}{(9.1 \times 10^{-31}) \times 10^7} = \mathbf{7.3 \times 10^{-11}\ m}$$

This wavelength is comparable with the atomic dimensions. Hence a **moving electron exhibits a wave character.**

EXAMPLE 2. *Show that the de Broglie wavelength associated with an electron of energy V electron-volts is approximately* $(1.227/\sqrt{V}\,)$ *nm.*

SOL. The de Broglie wavelength λ associated with an electron of mass m and energy E is given by

$$\lambda = \frac{h}{\sqrt{(2mE)}}$$

Here, kinetic energy $E_k = V\,\text{eV} = 1.6 \times 10^{-19}\ V\,\text{J}$

$$\lambda = \frac{6.62 \times 10^{-34}}{\sqrt{(2 \times 9.1 \times 10^{-31} \times 1.6 \times 10^{-19}\ V)}}$$

$$= \frac{1.227 \times 10^{-9}\,\text{m}}{\sqrt{V}} = \frac{1.227}{\sqrt{V}}\,\mathbf{nm}$$

EXAMPLE 3. *Find the kinetic energy of a proton whose de Broglie wavelength is 1 fm.*

SOL. $$pc = (mv)\,c = hc/\lambda$$

$$= \frac{(4.136 \times 10^{-15}\ \text{eV.s})\ (3 \times 10^8\ \text{ms}^{-1})}{1 \times 10^{-15}\,\text{m}} = 1.241\ \text{GeV.}$$

Rest energy of proton = E_0 = 0.938 GeV.

$pc > E_0$. Hence a relativistic calculation is needed.

The total energy of the proton is

$$E = \sqrt{E_0^2 + p^2c^2} = \sqrt{(0.938\ GeV)^2 + (1.241\ GeV)^2} = 1.556\ \text{GeV.}$$

The kinetic energy of the proton is

$$KE = E - E_0 = (1.556 - 0.938)\ \text{GeV} = \mathbf{0.618\ GeV}$$

EXAMPLE 4. *Show that the de Broglie wavelength for a material particle of rest mass* m_0 *and charge q, accelerated from rest through a potential difference of V volts relativistically is given by*

$$\lambda = \frac{h}{\sqrt{2m_0 qV\left(1 + \dfrac{qV}{2m_0c^2}\right)}}$$

SOL. We use the relativistic formula to find momentum.

Kinetic energy $\qquad E_k = qV.$

$$E^2 = p^2c^2 + m_0^2c^4,\ E = E_k + m_0c^2 = qV + m_0c^2$$

So, $$p^2c^2 = E^2 - m_0^2c^4 = (qV + m_0c^2)^2 - m_0^2c^4 = q^2V^2 + 2m_0c^2qV$$

$$p^2 = 2m_0\,qV\left(1+\frac{qV}{2m_0\,c^2}\right) \text{ or } p = \sqrt{2m_0\,qV\left(1+\frac{qV}{2\,m_0\,c^2}\right)}$$

$$\lambda = \frac{h}{p} = \frac{h}{\sqrt{2\,m_0\,qV\left(1+\dfrac{qV}{2\,m_0\,c^2}\right)}}$$

Special Case. If the charged particle is an electron, then $q = e = 1.6 \times 10^{-19}$ C, $m_0 = 9.1 \times 10^{-31}$ kg.

$$\lambda = \frac{6.62\times10^{-34}}{\sqrt{(2\times9.1\times10^{-31}\times1.6\times10^{-19}V)\left(1+\dfrac{1.6\times10^{-19}\text{ V}}{2\times9.1\times10^{-31}\times(3\times10^{8})^2}\right)}}$$

$$= \frac{1.227}{\sqrt{\ }}\times\frac{1}{\sqrt{1+9.768\times10}}\text{ nm}$$

Phase velocity (or Wave Velocity) of de Broglie Waves

A particle of mass m moving with velocity v has a wave associated with it whose wavelength is given by

$$\lambda = \frac{h}{mv}\cdot$$

Let E be the total energy of the particle. Let ν be the frequency of the associated wave. We equate the quantum expression $E = h\nu$ with the relativistic formula for total energy $E = mc^2$. So we get

$$h\nu = mc^2 \text{ or } \nu = mc^2/h.$$

Let v_p be the de Broglie wave velocity. Then,

$$v_p = \nu\lambda = \left(\frac{mc^2}{h}\right)\left(\frac{h}{mv}\right) = \frac{c^2}{v}\cdot$$

But the particle velocity v is always less than c (the velocity of light). Therefore, the de Broglie wave velocity v_p must be greater than c.

11.2 Expression for Group Velocity

Consider two waves that have the same amplitude A but differ by an amount $\Delta\omega$ in angular frequency and an amount Δk in wave number. They can be represented by the equations

$$y_1 = A\cos(\omega t - kx)$$
$$y_2 = A\cos[(\omega+\Delta\omega)\,t - (k+\Delta k)\,x]$$

The superpostition of the two waves will yield a single *wave packet* or *wave group.* Let us find the velocity v_g with which the wave group travels.

The resultant displacement y at any time t and any position x is the sum of y_1 and y_2.

$$\begin{aligned} y &= y_1 + y_2 \\ &= A\cos(\omega t - kx) + A\cos[(\omega+\Delta\omega)\,t - (k+\Delta k)\,x] \\ &= 2A\cos\frac{1}{2}[2\omega+\Delta\omega]t-(2k+\Delta k)x]\cos\frac{1}{2}(\Delta\omega t - \Delta kx) \end{aligned}$$

$\Delta\omega$ and Δk are small compared with ω and k respectively. Therefore,

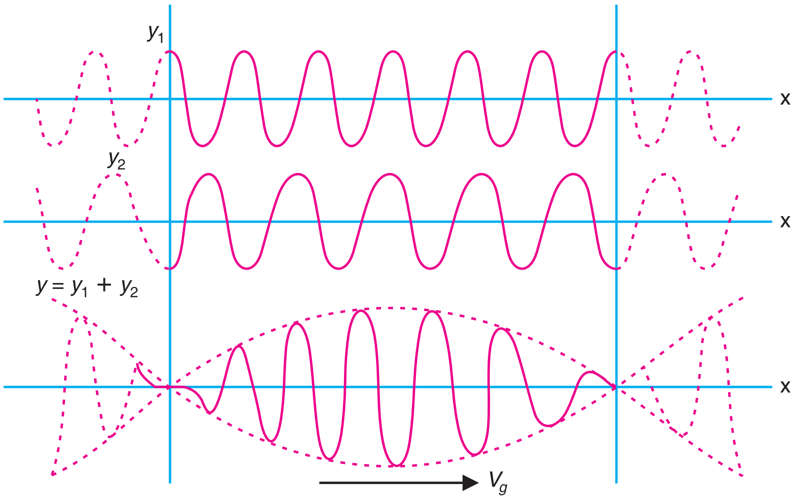


Fig. 11.1

$$2\omega + \Delta\omega \approx 2\omega$$

$$2k + \Delta k \approx 2k.$$

$$\therefore \quad y = 2A\cos\left(\frac{\Delta\omega}{2}t - \frac{\Delta k}{2}x\right)\cos(\omega t - kx) \quad ...(1)$$

This is the analytical expression for resultant wave (wave packet) due to superposition of the two waves. The second cosine function is the original wave. The coefficient of this cosine can be considered to be an amplitude that varies with x and t. This variation of amplitude is called the *modulation* of the wave.

Hence Eq. (1) represents a wave of angular frequency ω and wave number k that has superimposed upon it a modulation of angular frequency $\frac{1}{2}\Delta\omega$ (Fig. 11.1).

The velocity v_g of the wave groups is

$$v_g = \frac{\Delta\omega}{\Delta k} \quad ...(2)$$

When ω and k have continuous spreads, the group velocity is given by

$$v_g = \frac{d\omega}{dk} \quad ...(3)$$

This is the expression for the group velocity.

Group Velocity of de Broglie Waves

A particle moving with a velocity v is supposed to consist of a group of waves, according to de Broglie hypothesis.

The group velocity is given by,

$$v_g = \frac{d\omega}{dk}$$

The angular frequency and wave number of the de Broglie waves associated with a particle of rest mass m_0 moving with the velocity v are given by

$$\omega = 2\pi\nu = \frac{2\pi mc^2}{h} = \frac{2\pi m_0 c^2}{h\sqrt{1-v^2/c^2}} \quad ...(1)$$

and $$k = \frac{2\pi}{\lambda} = \frac{2\pi mv}{h} = \frac{2\pi m_0 v}{h\sqrt{1-v^2/c^2}} \qquad ...(2)$$

By differentiation, we obtain

$$\frac{d\omega}{dv} = \frac{2\pi m_0 v}{h(1-v^2/c^2)^{3/2}}$$

$$\frac{dk}{dv} = \frac{2\pi m_0}{h(1-v^2/c^2)^{3/2}}\,.$$

The group velocity v_g of the de Broglie waves associated with particle is

$$v_g = \frac{d\omega}{dk} = \frac{d\omega/dv}{dk/dv} = v$$

Hence the de Broglie wave group associated with a moving particle travels with the same velocity as the particle.

Relation between Group Velocity (v_g) and Wave Velocity or Phase Velocity (v_p)

We have the relations,

Wave velocity, $$v_p = \frac{\omega}{k} \qquad ...(1)$$

Group velocity, $$v_g = \frac{d\omega}{dk} \qquad ...(2)$$

The wave number is given by

$$k = \frac{2\pi}{\lambda}$$

∴ $$\frac{dk}{d\lambda} = -\frac{2\pi}{\lambda^2} \qquad ...(3)$$

Also $$\omega = 2\pi\nu = 2\pi\frac{v_p}{\lambda}$$

∴ $$\frac{d\omega}{d\lambda} = 2\pi\left[-\frac{v_p}{\lambda^2} + \frac{1}{\lambda}\frac{dv_p}{d\lambda}\right]$$

or $$\frac{d\omega}{d\lambda} = -\frac{2\pi}{\lambda^2}\left[v_p - \lambda\frac{dv_p}{d\lambda}\right] \qquad ...(4)$$

Dividing Eq. (4) by Eq. (3), we get

$$\frac{d\omega}{d\lambda}\cdot\frac{d\lambda}{dk} = \frac{-\frac{2\pi}{\lambda^2}\left[v_p - \lambda\frac{dv_p}{d\lambda}\right]}{-\frac{2\pi}{\lambda^2}}$$

or $$\frac{d\omega}{dk} = v_p - \lambda\frac{dv_p}{d\lambda}$$

∴ $$v_g = v_p - \lambda\frac{dv_p}{d\lambda} \qquad ...(5)$$

Eq. (5) gives the relationship between group velocity (v_g) and phase velocity or wave velocity (v_p).

From this equation the following two cases arise.

(*i*) *For dispersive medium*

$v_p = f(\lambda)$. Usually $dv_p/d\lambda$ is positive (normal dispersion).

$\therefore \qquad v_g < v_p.$

This is the case with de Broglie waves.

(*ii*) *For non-dispersive medium*

$$v_p \neq f(\lambda). \ \frac{dv_p}{d\lambda} = 0. \quad \therefore v_g = v_p$$

This result is true for electromagnetic waves in vacuum.

11.3 Experimental Study of Matter Waves

Davisson and Germer's experiment.

Experimental arrangement. Fig. 11.2 shows the experimental arrangement. Electrons are produced by heating a filament (F) by a low tension battery (L.T). The electrons are then accelerated through a potential difference V in the electron gun. This electron beam falls on a large single crystal of nickel. The electrons are scattered in all directions by the atoms in the crystal. The detector can be moved to any angle ϕ relative to the incident beam. The energy of the electrons in the primary beam, the angle at which they reach the target, and the position of the detector could all be varied.

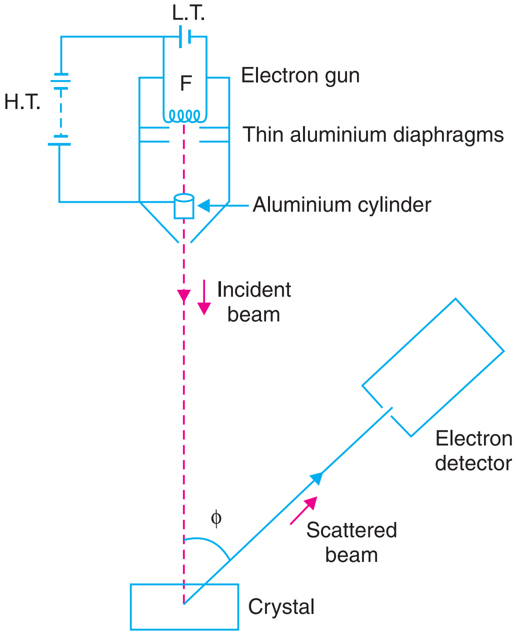


Fig. 11.2

Experimental procedure. The accelerating potential V is given a low value. The beam of electrons falls normally on the surface of the crystal. The detector is moved to various positions and the intensity of the diffracted beam at each position is noted. The intensity of diffracted beam is plotted against the angle between the incident beam and the beam entering the detector. The method of plotting is such that the intensity at any angle is proportional to the distance of the curve at that angle from the point of scattering. The observations are repeated for different accelerating voltages and a number of curves are drawn. Fig.11.3(*i*) shows typical polar graphs of electron intensity.

(*i*) The graph remains fairly smooth, till the accelerating voltage becomes 44*V* when a '***bump***' appears on the curve.

(*ii*) As the accelerating voltage is increased, the length of the '***bump***' increases.

(*iii*) The '***bump***' becomes most prominent in the curve for 54 volt electrons at $\phi = 50°$.

(*iv*) As the accelerating voltage is further increased, the bump decreases in length and finally disappears at 68*V*.

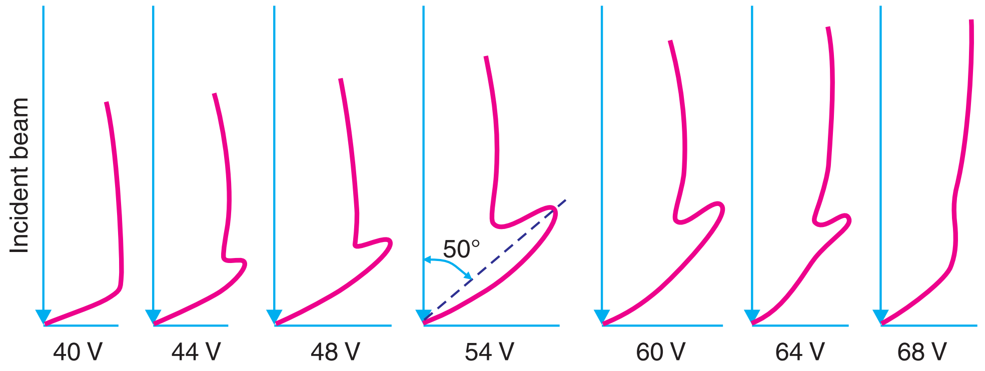


Fig. 11.3(*i*)

Calculation of Observed Wavelength

For nickel, the spacing of the atomic planes which can be measured by x-ray diffraction, is $d = 0.091$ nm.

There is an intense reflection of the beam at the angle $\phi = 50°$.

The angle of incidence relative to the family of Bragg planes shown in Fig.11.3 (*ii*) is

$$\theta = \frac{180 - 50}{2} = 65°.$$

The Bragg equation for maxima in the diffraction pattern is

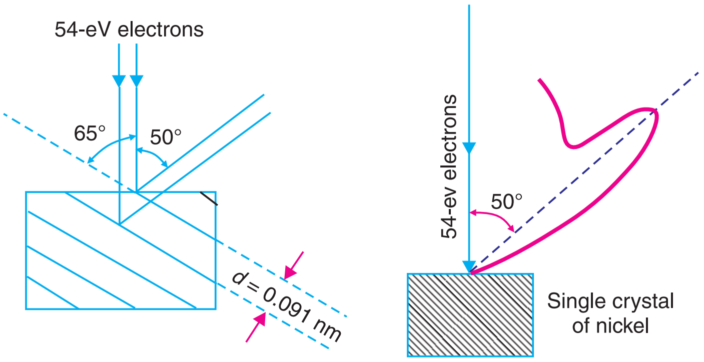


Fig. 11.3(*ii*)

$$n\lambda = 2d \sin\theta \qquad ...(i)$$

For $n = 1$ the de Broglie wavelength λ of the diffracted electrons is

$$\lambda = 2d \sin\theta = (2)(0.091 \text{ nm})(\sin 65°) = 0.165 \text{ nm}$$

Calculation of the expected wavelength of the electrons using de Broglie's formula. The electron kinetic energy KE is 54 eV. The electron wavelength is therefore

$$\lambda = \frac{h}{mv} = \frac{h}{\sqrt{2mKE}} = \frac{6.63\times10^{-34}\ J.s}{\sqrt{(2)(9.1\times10^{-31}\ kg)(54\ eV)(1.6\times10^{-19}\ J/eV)}} = 0.166 \text{ nm}$$

The observed wavelength agrees well with the expected wavelength. The Davisson-Germer experiment thus directly verifies de Broglie's hypothesis of the wave nature of moving bodies.

G.P. Thomson's Experiment

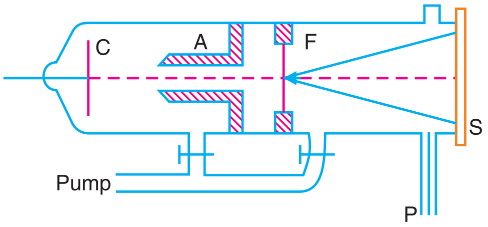


Fig. 11.4 (*i*)

Experimental arrangement. The experimental arrangement is shown in Fig. 11.4 (*i*). A beam of cathode rays is produced in a discharge tube *AC* by means of an induction coil. The electrons passing through a fine hole *A*, are incident on a thin gold foil *F*. The thickness of the foil is about 10^{-8} m. The emergent beam of electrons is received on a photographic plate *P*. Visual examination of the pattern is made possible by the fluorescent screen *S*. A very high vacuum is maintained in the camera part *FP* of the apparatus while air is allowed to leak into the discharge tube section through a needle valve. Thus a beam of required voltage can be continuously produced, the discharge tube being sufficiently soft.

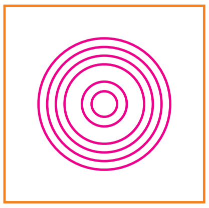


Fig. 11.4 (*ii*)

Experimental procedure. A beam of electrons of known velocity is made to fall on the photographic plate, after traversing the thin gold foil. When the plate is developed a symmetrical pattern consisting of concentric rings about a central spot, is obtained [Fig. 11.4 (*ii*)]. This pattern is similar to the pattern produced by X-rays in the powdered crystal method. When the cathode rays in the discharge tube are deflected by a magnetic field, the entire pattern on the screen *S* is found to shift correspondingly. Thus the pattern is confirmed as due to diffracted electrons and not due to secondary X-rays, generated by the electrons going through the foil. Further, on removing the film *F* the pattern diappears, showing that the presence of the film is essential. If the electrons behaved as corpuscles, the electrons passing through the foil should have been scattered through a wide angle. Clearly, this experiment demonstrates that the electron beam behaves as waves, since diffraction patterns can be produced only by waves.

Verification of the de Broglie equation. G.P. Thomson employed very high voltages of the order of 50000 volts to accelerate the electrons. With very high speed electrons, relativistic correction for the mass of the electron has to be applied. It can be shown that

$$\lambda = \frac{12.27}{\sqrt{V}} \frac{1}{\sqrt{1 + 9.836 \times 10^{-7}\ V}}\ \text{Å}. \qquad ...(3)$$

The correction factor is very small, except for very large values of *V*.

To calculate λ from the radii of the rings. In the polycrystalline film there will be some crystals set at the correct angle to give a Bragg reflection. If there are enough crystals distributed at random, the result of such reflections will be a series of rings, arising from the intersection of the cones of diffraction with the photographic plate. Let *AB* be the incident beam passing though the film at *B*. *BP* is the direction of the beam which has suffered a Bragg reflection in some crystal in the film at *B*. This reflected beam falls at the point *P* on the photographic plate at a distance *R* from the central point *O* (Fig. 11.5). Let the distance *BO*, from the film to the plate, be *L*. $\angle PBO = 2\,\theta$ where θ is the glancing angle given by the Bragg relation,

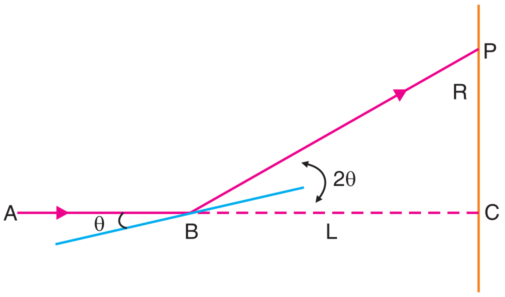


Fig. 11.5

$$n\lambda = 2d \sin\theta.$$

$R/L = \tan 2\theta = 2\theta$, since θ is small.

$$\therefore \qquad \theta = R/2L$$

But $2d \sin\theta = 2d\theta = n\lambda$ or $\theta = n\lambda/2d$

$$\therefore \qquad \frac{n\lambda}{2d} = \frac{R}{2L}$$

$$\text{or} \qquad \lambda = \frac{Rd}{nL} \qquad ...(2)$$

From this the wavelength is calculated. This agrees with the calculated value given by Eqn. (1). This provides ultimate confirmation for the wave nature of the electron.

Example 1. 10 *kV electrons are passed through a thin film of a metal for which the atomic spacing is* 5.5 × 10^{-11} *m. What is the angle of deviation of the first order diffraction maximum?*

Sol. Wavelength of the electron = $\lambda = \dfrac{h}{\sqrt{mVe}}$

$$\therefore \qquad \lambda = \frac{6.626 \times 10^{-34}}{\sqrt{2\,(9.11 \times 10^{-31})\,10^{4}\,(1.602 \times 10^{-19})}}$$

$$= 1.227 \times 10^{-11} \text{ m.}$$

Applying Bragg's formula for diffraction at the atomic planes,

$$n\lambda = 2d \sin\theta \text{ or } 1 \times (1.227 \times 10^{-11}) = 2 \times (5.5 \times 10^{-11}) \sin\theta$$

or $\quad \sin\theta = 0.1115. \; \therefore \; \theta = 6°24'.$

Angle through which electron is deviated = $2\theta = 12°48'$.

Two slit interference pattern with electrons

The interference of electron waves can be exhibited by a *double slit experiment.* The interference pattern so obtained resembles the pattern in case of visible light and proves the association of wave-packets with electrons.

The experimental arrangement is shown in Fig. 11.6.

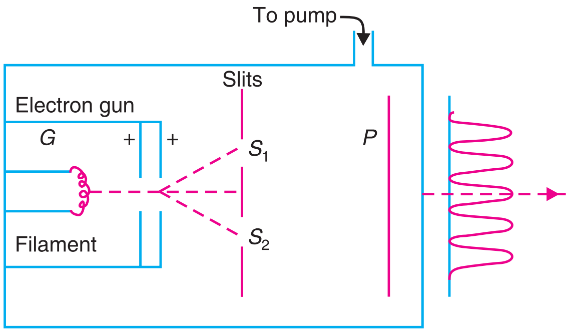


Fig. 11.6

The electron gun (G) supplies a mono-energetic beam of electrons. The electrons emerging from the gun are allowed to pass through two slits S_1 and S_2. The interference pattern is observed on a photographic film P. The entire apparatus is enclosed in a high vacuum chamber, so that electrons emerging from the gun travel to the screen without collisions in the path. When photographic film is observed with an electron microscope, it is observed that interference fringes are obtained on the film. The distribution of intensity on the film is shown on the right. The formation of interference fringes clearly indicates the wave nature of electrons.

Knowing the separation between S_1, S_2 and the distance between the plane of S_1, S_2 and the photographic plate P, the value of wavelengh associated with electrons may be calculated. This value comes out to be same as given by de-Broglie relation.

11.4 Heisenberg's Uncertainty Principle

Statement. *It is impossible to determine precisely and simultaneously the values of both the members of a pair of physical variables which describe the motion of an atomic system.*

Such pairs of variables are called *canonically conjugate* variables.

Example. According to this principle, the position and momentum of a particle (say electron) cannot be determined simultaneously to any desired degree of accuracy.

Taking Δx as the error in determining its position and Δp the error in determining its momentum at the same instant, these quantities are related as follows:

Werner Heisenberg
(1901 – 1976)

$$\Delta x \, \Delta p = h/2\pi.$$

The product of the two errors is approximately of the order of Planck's constant. If Δx is small, Δp will be large and vice versa. It means that if one quantity is measured accurately, the other quantity becomes less accurate. Thus any instrument cannot measure the quantities more accurately than predicted by Heisenberg's principle of uncertainty or indeterminacy. The same relation holds for the energy and time related to any given event.

i.e., $$\Delta E \, \Delta t = h/2\pi$$

According to classical ideas, it is possible for a particle to occupy a fixed position and have a definite momentum and we can predict *exactly* its position and momentum at any time later. But according to the uncertainty principle, it is not possible to determine accurately the simultaneous values of position and momentum of a particle at any time. Heisenberg's principle implies that in physical measurements probability takes the place of exactness and as such phenomena which are impossible according to classical ideas may find a small but finite probability of occurrence.

Illustration (i) : Determination of position with γ-ray-microscope. Suppose we try to measure the position and linear momentum of an electron using an imaginary microscope with a very high resolving power (Fig. 11.7). The electron can be observed if atleast one photon is scattered by it into the microscope lens. The resolving power of the microscope is given by the relation

$$\Delta x = \frac{\lambda}{2\sin\theta}$$

where Δx is the distance between two points which can be just resolved by microscope. This is the range in which the electron would be visible when disturbed by the photon. Hence Δx is the uncertainty involved in the position measurement of the electron.

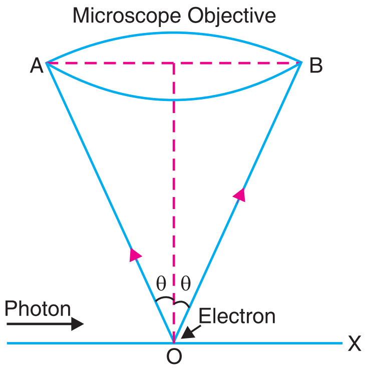


Fig. 11.7

However, the incoming photon will interact with the electron through the Compton effect. To be able to see this electron, the scattered photon should enter the microscope within the angle 2θ. The momentum imparted by the photon to the electron during the impact is of the order of h/λ. The component of this momentum along OA is $-\frac{h}{\lambda}\sin\theta$ and that along OB is $\frac{h}{\lambda}\sin\theta$.

Hence the uncertainty in the momentum measurement in the x-direction is

$$\Delta p_x = \frac{h}{\lambda}\sin\theta - \left(-\frac{h}{\lambda}\sin\theta\right) = \frac{2h}{\lambda}\sin\theta.$$

$$\therefore \qquad \Delta x \times \Delta p_x = \frac{\lambda}{2\sin\theta} \times \frac{2h}{\lambda}\sin\theta = h.$$

A more sophisticated approach will show that $\Delta x \, \Delta p_x \geq h/2\pi$.

It is clear that the process of measurement itself perturbs the particle whose properties are being measured.

Illustration (ii) : Diffraction of a beam of electrons by a slit. A beam of electrons is transmitted through a slit and received on a photographic plate P kept at some distance from the slit (Fig 11.8). We can only say that the electron must have passed through the slit and cannot specify its exact location in the slit as the electron crosses it. Hence the position of any electron recorded on the plate is uncertain by an amount equal to the width of the slit (Δy). Let λ be the wavelength of the electrons and θ be the angle of deviation corresponding to first minimum. From the theory of diffraction in optics, $\Delta y = \dfrac{\lambda}{\sin\theta}$. This is the uncertainty in determining the position of electron along y-axis.

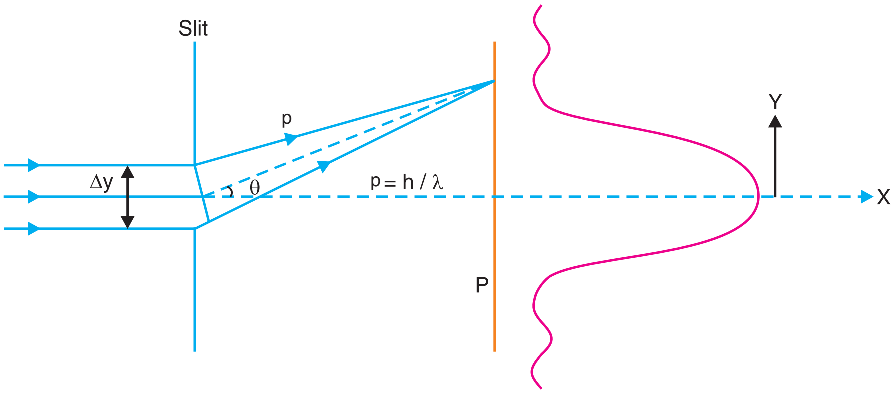


Fig. 11.8

Initially the electrons are moving along X-axis and so they have no component of momentum along y-axis. As the electrons are deviated at the slit from their initial path to form the pattern, they acquire an additional component of momentum along y-axis. If p is the momentum of the electron on emerging from the slit, the component of momentum of electron along y-axis is $p \sin\theta$. As the electron may be anywhere within the pattern from angle $-\theta$ to $+\theta$, the y-component of momentum of the electron may be anywhere between $-p\sin\theta$ and $+p\sin\theta$.

Therefore, the uncertainty in the y-component of momentum of the electron

$$\Delta p_y = 2p\sin\theta = \frac{2h}{\lambda}\sin\theta \qquad \left(\text{since } \lambda = \frac{h}{p}\right)$$

$$\therefore \qquad \Delta y\,\Delta p_y = \frac{\lambda}{\sin\theta}\times\frac{2h}{\lambda}\sin\theta = 2h.$$

i.e., $\Delta y\,\Delta p_y \geq h/2\pi$, which is Heisenberg's uncertainty principle.

EXAMPLE 1. *A microscope, using photons, is employed to locate an electron in an atom to within a distance of* 0.2 Å. *What is the uncertainty in the momentum of the electron located in this way?*

SOL. Here, $\Delta x = 0.2\ \text{Å} = 0.2\times10^{-10}\ m$. $\Delta p = ?$

We have, $\qquad \Delta x\,\Delta p \approx \dfrac{h}{2\pi}$ or $\Delta p = \dfrac{h}{2\pi\Delta x}$

$$\therefore \qquad \Delta p = \frac{6.626\times10^{-34}}{2\pi(0.2\times10^{-10})} = 5.274\times10^{-24}\ \text{kg ms}^{-1}.$$

EXAMPLE 2. *Prove the nonexistence of electron in the nucleus on the basis of uncertainty principle.*

SOL. Nuclear radius (r_0) is of the order of 10^{-14} m. If the electron is confined within nucleus, the uncertainty in its position is

$$\Delta x = 2r_0 = 2\times10^{-14}\ \text{m}.$$

The uncertainty in the electron momentum is

$$\Delta p \cong \frac{\hbar}{2r_0} \cong \frac{1.055\times10^{-34}}{2\times10^{-14}} \cong 5.28\times10^{-21}\ \text{kgms}^{-1}$$

The momentum of the electron must at least be of this order.

$$\text{K.E. of the electron} = T = \frac{p^2}{2m} = \frac{\left(5.28 \times 10^{-21}\right)^2}{2\left(9.1 \times 10^{-31}\right)} \text{ J} \cong 95.7 \text{ MeV}$$

However, electrons emitted during β-decay have energies of the order of 3 MeV, which varies widely from our calculated value. Hence we conclude that electrons do not reside in the nucleus.

Example 3. *The lifetime of an excited state of an atom is about* 10^{-8} *sec. Calculate the minimum uncertainty in the determination of the energy of the excited state.*

Sol. We have, $\Delta E \, \Delta t \geq h/2\pi$.

$$\therefore \qquad \Delta E \geq \frac{h}{2\pi \Delta t} = \frac{6.6 \times 10^{-34}}{2\pi \, (10^{-8})}$$

$$\therefore \qquad \Delta E \geq 1.0 \times 10^{-26} \, J = 6.5 \times 10^{-8} \, eV.$$

This is known as the *energy width* of an excited state.

Example 4. *Consider an electron of momentum p in the Coulomb field of a proton. The total energy is*

$$E = \frac{p^2}{2m} - \frac{e^2}{(4\pi\varepsilon_0) r},$$

where r is the distance of the electron from the proton. Assuming that the uncertainty Δr *of the radial coordinate is* $\Delta r \approx r$ *and that* $\Delta p \approx p$*, use Heisenberg's uncertainty principle to obtain an estimate of the size and the energy of the hydrogen atom in the ground state.*

Sol. From the uncertainty principle, $\Delta p \geq \dfrac{(h/2\pi)}{\Delta r}$.

Making the assumption $\Delta p \geq \dfrac{(h/2\pi)}{\Delta r} = \dfrac{(h/2\pi)}{r}$, we obtain

$$E = \frac{(h/2\pi)^2}{2m} \frac{1}{r^2} - \left(\frac{e^2}{4\pi\varepsilon_0} \frac{1}{r} \right)$$

The radius at which E is a minimum is given by the condition $dE/dr = 0$, from which we find

$$r = \frac{(4\pi\varepsilon_0)\,(h/2\pi)^2}{me^2} = a_0.$$

The corresponding lowest value of E is

$$E_0 = -\frac{e^4 m}{(4\pi\varepsilon_0)^2 \, 2(h/2\pi)^2} = -13.6 \, eV.$$

Complementarity Principle of Bohr. *In any experimental situation in which a physical entity (matter or radiation) exhibits its wave properties, it is impossible to attribute corpuscular characteristics to it. In other words, the particle and wave aspects of a physical entity are complementary and cannot be exhibited at the same time.*

For example, in a two slit interference experiment if we try to define electron trajectory either by closing one slit or by placing some detecting apparatus immediately behind one of the slits, the interference pattern disappears. It follows that any experiment which can be devised either displays wave like or particle like characteristics of the system. The wave and particle pictures give complementary descriptions of the same system.

11.5 Wave Mechanical Atom Model

The most recent development in the theory of atomic structure shows that the Bohr picture of the atoms with sharply defined electron orbits is not correct. According to the Wave mechanics developed by Heisenberg, Schrodinger and Dirac, the electrons move around the nucleus as *wave packets*, which are formed in a somewhat similar manner as standing waves are formed in sound.

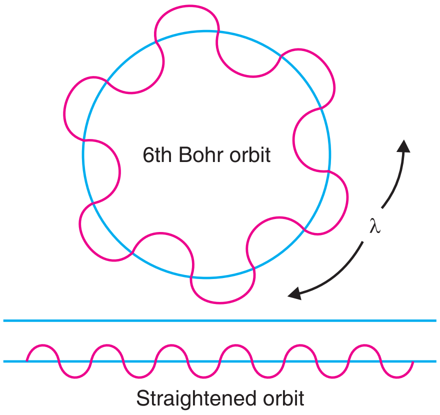


Fig. 11.9

This atom model has electrons in various orbits and the electrons behave as matter waves of wavelength $\lambda = h/mv$. The electron exists as a standing wave in each orbit. The energy level and 'orbits' of Bohr model are retained. Even the quantum conditions of Bohr follow from wave mechanics. For example, consider the electron in the hydrogen atom as a *standing wave* extending in a circle round the nucleus. In order that this wave may just occupy the circumference of a circle, the circle must contain an integral number of wavelengths (Fig. 11.9). *i.e.*, $2\pi r = n\lambda$ where r is the radius of the circle and n, an integer.

But $$\lambda = \frac{h}{mv}.$$

$$\therefore \qquad 2\,\pi r = n\frac{h}{mv} \quad \text{or} \quad mvr = n\frac{h}{2\pi}.$$

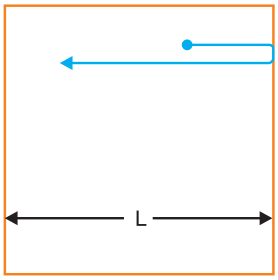


Fig. 11.10

But mvr is the angular momentum of the electron, regarded as a particle. Hence the angular momentum is equal to an integral multiple of $h/2\pi$ (Bohr's postulate).

The particle in a box

Let us discuss the possible energy states of a particle in a box on the basis of de Broglie's hypothesis. Consider a particle of mass m enclosed in a box of length L with impenetrable walls (Fig. 11.10). Suppose the (bound) particle is moving back and forth in the x-direction with constant speed v, making perfectly elastic collisions with the walls of the box. Since the walls of the box are impenetrable, the particle cannot move beyond the walls and so the amplitude of the associated wave must drop to zero at the walls. In other words, the particle moving back and forth between opposite walls will form a stationary-wave pattern with nodes at the walls. Wave functions of the particle trapped in the box are shown in Fig. 11.11. The general formula for the permitted de Broglie wavelengths of the particle is

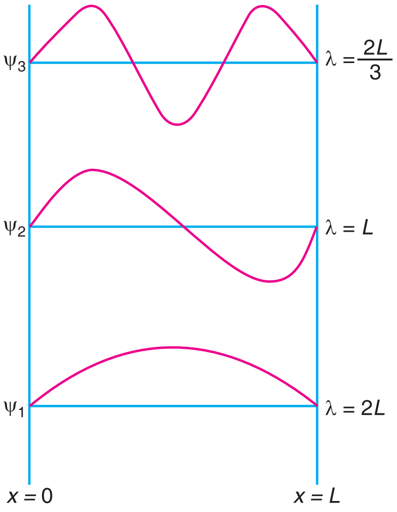


Fig. 11.11

$$\lambda_n = \frac{2L}{n},\; n = 1, 2, 3 \qquad \text{...(1)}$$

The possible values of the momentum of the particle are therefore,

$$p = \frac{h}{\lambda_n} = n\frac{h}{2L}.$$

Consequently, the possible values of the kinetic energy of the particle are

$$E_n = \frac{p^2}{2m} = n^2 \frac{h^2}{8mL^2} \qquad ...(2)$$

Thus only certain discrete energy states are possible for the particle bound in a box. Each permitted energy is called an *energy level.* The integer n that specifies an energy level E_n is called its *quantum number.*

11.6 Mathematical Proof of Uncertainty Principle for One Dimensional Wave-packet

We shall derive the position-momentum uncertainty relation by using the theory of Fourier analysis. A moving particle corresponds to a single wave group. An isolated wave group is the result of superposing an infinite number of waves with different angular frequencies ω, *continuous* range of wave numbers k and amplitudes (Fig. 11.12). The composition produces oscillations confined to a *single* region of space and thus provides an idealized picture of a localized matter wave.

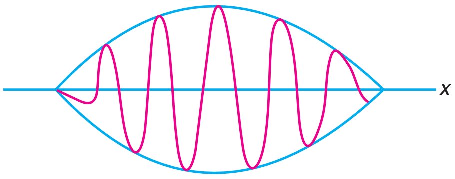


Fig. 11.12

At a certain time t, the wave group ψ (x) can be represented by the *Fourier integral*

$$\psi(x) = \int_0^\infty g\,(k)\cos kx\,dk$$

Here the amplitude function g (k) describes how the amplitudes of the waves that contribute to ψ (x) vary with wave number k. ψ (x) and g (k) are just *Fourier transforms* of each other. Fig. 11.13 shows Gaussian distributions for the amplitude function g (k) and the wave packet ψ (x). The relationship between the distance Δ x and the wave number spread Δ k depends upon the shape of the wave group and upon how Δ x and Δ k are defined. The widths Δ x and Δ k obey a reciprocal relation in which the product Δx Δk is equal to pure number. The minimum value of the product Δ x Δ k occurs when the envelope of the group has the bell shape of a Gaussian function (Fig. 11.13). Thus, the Gaussian wave packets happen to be minimum uncertainty wave packets. If Δ x and Δ k are taken as the standard deviations of the respective functions ψ (x) and g (k), then this minimum value is 1/2. Wave groups in general do not have Gaussian forms. So we can write

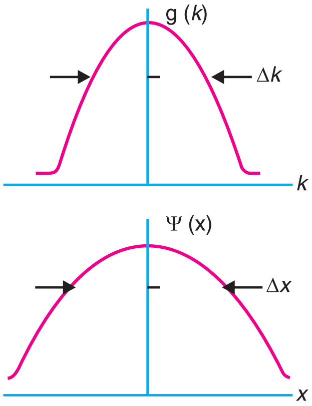


Fig. 11.13

$$\Delta x\,\Delta k \geq \frac{1}{2} \qquad ...(1)$$

Let λ be the de Broglie wavelength of the particle. We see from

$$k = \frac{2\pi}{\lambda} = \frac{2\pi p}{h}$$

that the momentum of the particle is determined by the wave number k.

$$\therefore \qquad p = \frac{hk}{2\pi}\cdot$$

$$\Delta p = \frac{h\,\Delta k}{2\pi} = \hbar\,\Delta k\,\cdot$$

Hence an uncertainty $\Delta\, k$ in the wave number of the de Broglie waves associated with the particle results in an uncertainty $\Delta\, p$ in the particle's momentum.

From Eq. (1), $\Delta\, x\, \Delta\, k \geq 1/2$ or $\Delta\, k \geq 1/2\, \Delta\, x$

$$\therefore \qquad \Delta x\, \Delta p \;\geq\; \hbar/2 \qquad \text{...(2)}$$

This is the *Heisenberg uncertainty relation for position and momentum*, according to which the uncertainty Δx in measuring the x coordinate of a particle is related to the uncertainty $\Delta\, p_x$ in measuring the x component of the momentum, the product of the uncertainties being equal to or greater than $\hbar/2$.

The three-dimensional form of the Heisenberg uncertainty relations for position and momentum is now

$$\Delta x\, \Delta p_x \;\geq\; \frac{\hbar}{2},\; \Delta\, y\, \Delta\, p_y \geq \frac{\hbar}{2},\; \Delta\, z\, \Delta\, p_z \geq \frac{\hbar}{2} \qquad \text{...(3)}$$

The theory of Fourier analysis may also be invoked to obtain a *time-energy uncertainty relation.* Indeed, according to Fourier analysis, a wave packet of duration $\Delta\, t$ must be composed of plane-wave components whose angular frequencies extend over a range $\Delta\omega$ such that $\Delta\, t\, \Delta\, \omega \geq 1/2$. Since $E = \hbar\omega$, we, therefore, have

$$\Delta t\, \Delta E \;\geq\; \hbar/2 \qquad \text{...(4)}$$

which is the *Heisenberg uncertainty relation for time and energy.* It connects the uncertainty ΔE in the determination of the energy of a system with the time interval $\Delta\, t$ available for this energy determination. Thus, if a system does not stay longer than a time $\Delta\, t$ in a given state of motion, its energy in that state will be uncertain by an amount $\Delta\, E \geq \hbar\,/\,2\,\Delta t$.

WAVE MECHANICS

11.7 Basic Postulates of Wave Mechanics

In the development of Wave Mechanics, there are certain basic postulates, which are of fundamental importance. The fundamental postulates are three in number. Other wave properties follow from them.

(1) *Each dynamical variable relating to the motion of a particle can be represented by a linear operator.*

Explanation. In classical Physics, certain definite functions of suitable variables are associated with each observable quantity. Thus (x,y,z) are associated with position, mv is associated with momentum, $\frac{1}{2}mv^2$ is associated with K.E. and so on. Similarly, in wave mechanics and quantum mechanics, certain operators are associated with observable quantities. For the x-component of the linear momentum of a particle which has a classical expression $p_x = m\left(\frac{dx}{dt}\right)$ we have a quantum mechanical operator $-i\left(\frac{h}{2\pi}\right)\frac{\partial}{\partial\, x}$. In the vector form, this operator is $-i\left(\frac{h}{2\pi}\right)\nabla$. For angular mementum we can write the operator as $(\mathbf{r} \times \mathbf{p}) = -i\left(\frac{h}{2\pi}\right)(\boldsymbol{r} \times \nabla)$. Similarly, for the observable *total energy*, the classical representation is $\frac{1}{2m}\,(p_x^{\,2} + p_y^{\,2} + p_z^{\,2}) + V\,(x, y, z)$ and the quantum mechanical operator is $-\frac{(h\,/\,2\,\pi)^2}{2\,m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) + V\;(x, y, z)$.

An operator tells us what operation to carry out on the quantity that follows it. The operator $i\left(\frac{h}{2\pi}\right)\frac{\partial}{\partial t}$ instructs us to take the partial derivative of what comes after it with respect to t and multiply the result by $i\left(\frac{h}{2\pi}\right)$.

Table 11.1 summarises the quantum operators for several physical quantities.

TABLE 11.1. Quantum operators

Quantity	*Classical definition*	*Quantum operator*
Position	$\mathbf{r}$	$\mathbf{r}$
Momentum	$\mathbf{p}$	$-i\frac{h}{2\pi}\nabla$
Angular momentum	$\mathbf{r}\times\mathbf{p}$	$-i\frac{h}{2\pi}\mathbf{r}\times\nabla$
Kinetic energy	$p^2/2m$	$-(h^2/8\pi^2 m)\,\nabla^2$
Total energy	$p^2/2m + E_p\,(\mathrm{r})$	$-(h^2/8\pi^2 m)\,\nabla^2 + E_p\,(r)$

(2) *A linear eigenvalue equation can be always linked with each operator.*

Example. The total energy operator is $i\left(\frac{h}{2\pi}\right)\frac{\partial}{\partial t}$. Consider the eigen value equation $i\left(\frac{h}{2\pi}\right)\frac{\partial\psi}{\partial t} = E\psi$. Here ψ is said to be an *eigenfunction* of the operator $i\left(\frac{h}{2\pi}\right)\frac{\partial}{\partial t}$ and E is called the corresponding energy *eigenvalue*.

(3) In general, when a measurement of a dynamical quantity a is made on a particle for which the wave function is ψ, we get *different* values of a during different trials. This is in conformity with the uncertainty principle. The most *probable value* of a is given by

$$<a> = \int_0^\infty \psi^* \hat{A}\psi \, dV$$

where $\hat{A}$ is the operator associated with the quantity a and ψ^* is the complex conjugate of ψ. The quantity $<a>$ is called the *expectation value* of $\hat{A}$ (that is the value of a obtained in the majority of the trials). The expectation value of momentum and energy may be found by using the corresponding differential operator. Thus

$$<\vec{p}> = \int_{-\infty}^{\infty} \psi^* \left(-\frac{ih}{2\pi}\nabla\right)\psi \, dx\, dy\, dz$$

$$<E> = \int_{-\infty}^{\infty} \psi^* \left(i\frac{h}{2\pi}\frac{\partial}{\partial t}\right)\psi \, dx\, dy\, dz\cdot$$

11.8 Derivation of Time-dependent form of Schrodinger Equation

The quantity that characterises the de Broglie waves is called the *wave function.* It is denoted by ψ. It may be a complex function. Let us assume that ψ is specified in the x direction by

$$\psi = Ae^{-i\omega(t - x/v)} \qquad \text{...(1)}$$

Erwin Schrodinger
(1887 – 1961)

If ν is the frequency, then $\omega = 2\pi\nu$ and $v = \nu\lambda$.

$$\therefore \qquad \psi = Ae^{-2\pi i(\nu t - x/\lambda)} \qquad \text{...(2)}$$

Let E be the total energy and p the momentum of the particle. Then $E = h\nu$ and $\lambda = h/p$. Making these substitutions in Eq. (2),

$$\psi = Ae^{-(2\pi i/h)(Et - px)} \qquad \text{...(3)}$$

Eq. (3) is a mathematical description of the wave equivalent of an unrestricted particle of total energy E and momentum p moving in the $+x$ direction.

Differentiating Eq. (3) twice with respect to x, we get

$$\frac{\partial^2\psi}{\partial x^2} = -\frac{4\pi^2 p^2}{h^2}\psi \qquad \text{...(4)}$$

Differentiating Eq. (3) once with respect to t, we get

$$\frac{\partial\psi}{\partial t} = -\frac{2\pi i E}{h}\psi \qquad \text{...(5)}$$

At speeds small compared with that of light, the total energy E of a particle is the sum of its kinetic energy $p^2/2m$ and its potential energy V. V is in general a function of position x and time t.

$$\therefore \qquad E = \frac{p^2}{2m} + V \qquad \text{... (6)}$$

Multiplying both sides of Eq. (6) by ψ we get

$$E\psi = \frac{p^2\psi}{2m} + V\psi \qquad \text{... (7)}$$

From Eqs. (5) and (4) we see that

$$E\psi = -\frac{h}{2\pi i}\frac{\partial\psi}{\partial t} \qquad \text{... (8)}$$

and

$$p^2\psi = -\frac{h^2}{4\pi^2}\frac{\partial^2\psi}{\partial x^2} \qquad \text{... (9)}$$

Substituting these expressions for $E\psi$ and $p^2\psi$ into Eq. (7) we obtain

$$-\frac{h}{2\pi i}\frac{\partial\psi}{\partial t} = -\frac{h^2}{8\pi^2 m}\frac{\partial^2\psi}{\partial x^2} + V\psi$$

or

$$\frac{ih}{2\pi}\frac{\partial\psi}{\partial t} = -\frac{h^2}{8\pi^2 m}\frac{\partial^2\psi}{\partial x^2} + V\psi \qquad \text{...(10)}$$

Eq. (10) is the *time-dependent form of Schrodinger's equation.*

In three dimensions the time-dependent form of Schrodinger's equation is

$$\frac{ih}{2\pi}\frac{\partial\psi}{\partial t} = -\frac{h^2}{8\pi^2 m}\left(\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2}\right) + V\psi \cdot$$

Schrodinger's equation : Steady-state form

In a great many situations the potential energy of a particle does not depend upon time explicitly. The forces that act upon it, and hence V, vary with the position of the particle only. When this is true, Schrodinger's equation may be simplified by removing all reference to t. The one-dimesional wave function ψ of an unrestricted particle may be written in the form

$$\psi = Ae^{-(2\pi i/h)(Et-px)}$$

$$= Ae^{-(2\pi iE/h)t}.e^{+(2\pi ip/h)x}$$

$$\therefore \qquad \psi = \psi_0\, e^{-(2\pi iE/h)t} \qquad \text{...(1)}$$

Here, $\psi_0 = Ae^{+(2\pi ip/h)x}$. That is, ψ is the product of a position dependent function ψ_0 and a time-dependent function $e^{-(2\pi iE/h)t}$

Differentiating Eq. (1) with respect to *t*, we get

$$\frac{\partial\psi}{\partial t} = -\frac{2\pi iE}{h}\psi_0\, e^{-(2\pi iE/h)t} \qquad \text{...(2)}$$

Differentiating Eq. (1) twice with respect to *x*, we get

$$\frac{\partial^2\psi}{\partial x^2} = \frac{\partial^2\psi_0}{\partial x^2}\, e^{-(2\pi iE/h)t} \qquad \text{...(3)}$$

We can substitute these values in the time-dependent form of Schrodinger's equation

$$\frac{ih}{2\pi}\frac{\partial\psi}{\partial t} = -\frac{h^2}{8\pi^2 m}\frac{\partial^2\psi}{\partial x^2} + V\psi$$

$$\therefore \qquad E\psi_0\, e^{-(2\pi iE/h)t} = -\frac{h^2}{8\pi^2 m}\frac{\partial^2\psi_0}{\partial x^2}e^{-(2\pi iE/h)t} + \mathrm{V}\psi_0\, e^{-(2\pi iE/h)t}$$

Dividing through by the common exponential factor, we get

$$\frac{\partial^2\psi_0}{\partial x^2} + \frac{8\pi^2 m}{h^2}(E-V)\,\psi_0 = 0 \qquad \text{...(4)}$$

Eq. (5) is the *steady-state form of Schrodinger's equation.*

In three dimensions it is

$$\nabla^2\psi_0 + \frac{8\pi^2 m}{h^2}(E-V)\,\psi_0 = 0 \qquad \text{...(5)}$$

Usually it is written in the form

$$\nabla^2\psi + \frac{8\pi^2 m}{h^2}(E-V)\,\psi = 0$$

11.9 Properties of the Wave Function

Physical significance of ψ. The probability that a particle will be found at a given place in space at a given instant of time is characterised by the function $\psi\,(x,y,z,t)$. It is called the wave function. This function can be either real or complex. The only quantity having a physical meaning is the square of its magnitude $P = |\psi|^2 = \psi\psi^*$ where ψ^* is the complex conjugate of ψ. The quantity *P* is the probability density. The probability of finding a particle in a volume *dx*, *dy dz* is $|\psi|^2\, dx\, dy\, dz$. Further, since the particle is certainly to be found somewhere in space

$$\iiint |\psi|^2\, dx\,.\,dy\,.\,dz = 1$$

the triple integral extending over all possible values of *x, y, z*.

A wave function (ψ) satisfying this relation is called a *normalised wave function.*

Orthogonal and normalised wave functions. If the product of a function $\psi_1\,(x)$ and the complex conjugate $\psi_2^*\,(x)$ of a function $\psi_2\,(x)$ vanishes when integrated with respect to *x* over the interval $a \le x \le b$, that is, if

$$\int_a^b \psi_2^*(x)\,\psi_1(x)\,dx = 0$$

then $\psi_1(x)$ and $\psi_2(x)$ are said to be orthogonal in the interval (a, b).

We know that the probability of finding a particle in the volume element dV is given by $\psi\psi^* \, dV$. The total probability of finding the particle in the entire space is, of course, unity, *i.e.*,

$$\int |\psi|^2 \, dV = 1.$$

where the integration extends over all space. The above equation can also be written as

$$\int \psi\psi^* \, dV = 1.$$

Any wave function satisfying the above equation is said to be normalised to unity or simply normalised.

Very often ψ is not a normalized wave function. We know that it is possible to multiply ψ by a constant A, to give a new wave function, $A\psi$, which is also a solution of the wave equation. Now the problem is to choose the proper value of A such that the new wave function is a normalized function. In order that it is a normalized function, it must meet the requirement

$$\int (\mathrm{A}\psi)^* \, \mathrm{A}\psi \, dx\,dy\,dz = 1$$

or, $$|A|^2 \int \psi\psi^* dx\,dy\,dz = 1$$

or $$|A|^2 = \frac{1}{\int \psi\psi^* \, dx\,dy\,dz}$$

$|A|$ is known as normalizing constant.

To arrive at results consistent with physical observations, several additional requirements are imposed on the wave function $\psi(x)$:

1. It must be well behaved, that is, single-valued and continuous everywhere.
2. If $\psi_1(x), \ldots \psi_n(x)$ are solutions of Schrodinger equation, then the linear combination $\psi(x) = a_1 \psi_1(x) + a_2 \psi_2(x) + \ldots a_n \psi_n(x)$ must be a solution.
3. The wave function $\psi(x)$ must approach zero as $x \to \pm\infty$

Eigenfunctions and Eigenvalues. Schrodinger's time-independent equation is an example of a type of differential equation called an *eigenvalue equation*. In general, we can write an eigenvalue equation as

$$F_{op}\,\psi = f\psi$$

The differential operator F_{op} operates on a function ψ, and this yields a constant f times the function. The function ψ is then called an *eigenfunction* of the operator F_{op}, and the corresponding value for f is called the *eigenvalue*.

SIMPLE APPLICATIONS OF SCHRODINGER'S EQUATION

11.10 The Particle in a Box : Infinite Square Well Potential

Consider a particle moving inside a box along the x-direction. The particle is bouncing back and forth between the walls of the box. The box has insurmountable potential barriers at $x = 0$ and $x = L$. *i.e.*, the box is supposed to have walls of infinite height at $x = 0$ and $x = L$ (Fig. 11.14). The particle has a mass m and its position x at any instant is given by $0 < x < L$.

The potential energy V of the particle is infinite on both sides of the box. The potential energy V of the particle can be assumed to be zero between $x = 0$ and $x = L$.

In terms of the boundary conditions imposed by the problem, the potential function is

$$V = 0 \text{ for } 0 < x < L$$

$$V = \infty \text{ for } x \le 0$$

$$V = \infty \text{ for } x \ge L$$

The particle cannot exist outside the box and so its wave function ψ is 0 for $x \le 0$ and $x \ge L$. Our task is to find what ψ is within the box, viz., between $x = 0$ and $x = L$.

Within the box, the Schrodinger's equation becomes

$$\frac{d^2\psi}{dx^2} + \frac{8\pi^2 m}{h^2} E\psi = 0$$

Putting $\frac{8\pi^2 mE}{h^2} = k^2$, the equation becomes

$$\frac{d^2\psi}{dx^2} + k^2\psi = 0$$

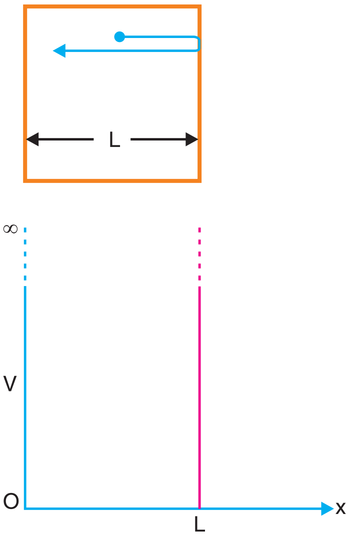


Fig. 11.14

The general solution of this equation is

$$\psi = A \sin kx + B \cos kx \qquad \text{...(1)}$$

The boundary conditions can be used to evaluate the constants A and B in equation (1).

$$\psi = 0 \text{ at } x = 0 \text{ and hence } B = 0$$

$$\psi = 0 \text{ at } x = L. \text{ Hence } 0 = A \sin kL$$

Since $A \ne 0$, $kL = n\pi$ where n is an integer or $k = \frac{n\pi}{L}$.

Thus
$$\psi_n(x) = A \sin \frac{n\pi x}{L} \qquad \text{...(2)}$$

$$\left.\begin{array}{r}\text{The energy of}\\ \text{the particle}\end{array}\right\} = E_n = \frac{k^2 h^2}{8\pi^2 m} = \frac{h^2 n^2 \pi^2}{L^2\, 8\pi^2 m}$$

$$\therefore \qquad E_n = \frac{n^2 h^2}{8mL^2} \qquad \text{...(3)}$$

For each value of n, there is an energy level and the corresponding wavefunction is given by equation (2). Each value of E_n is called an *eigenvalue* and the corresponding ψ_n is called *eigenfunction.* Thus inside the box, the particle can only have the *discrete* energy values specified by equation (3). Note also that the particle cannot have zero energy.

The particle in a box: Wave functions

It is certain that the particle is somewhere inside the box. Hence for a normalised wave function

$$\int_0^L \psi^* \psi \, dx = 1 \text{ } i.e., \text{ } A^2 \int_0^L \sin^2\left(\frac{n\pi x}{L}\right) dx = 1$$

i.e.,
$$A^2 \int_0^L \left(\frac{1 - \cos 2n\pi x/L}{2}\right) dx = 1 \text{ or } A^2 \frac{L}{2} = 1$$

or
$$A = \sqrt{\frac{2}{L}}$$

$$\therefore \left.\begin{array}{l}\text{The normalised wave}\\ \text{functions of the particle}\end{array}\right\} = \psi_n = \sqrt{\frac{2}{L}}\,\sin\frac{n\pi x}{L}$$

The normalised wave functions ψ_1 ψ_2 and ψ_3 are plotted in Fig. 11.15.

Example 1. *Calculate the permitted energy levels of an electron, in a box* 1 Å *wide.*

Sol. Here, m = mass of the electron $= 9.1 \times 10^{-31}$ kg;

$$L = 1\text{Å} = 10^{-10}\text{ m.}$$

$$E_n = ?$$

$$\therefore \text{ The permitted electron energies} = E_n = \frac{n^2h^2}{8mL^2}$$

$$= \frac{n^2\,(6.626\times10^{-34})^2}{8\,(9.1\times10^{-31})\,(10^{-10})^2}$$

$$= 6\times10^{-18}\,n^2\,J = 38n^2\,eV.$$

The minimum energy, the electron can have, is $E_1 = 38\ eV$, corresponding to $n = 1$.

The other values of energy are $E_2 = 4E_1 = 152\ eV$, $E_3 = 9E_1 = 342\ eV$ and so on.

Example 2. *A particle is moving in a one-dimensional box* (*of infinite height*) of width 10 Å. *Calculate the probability of finding the particle within an interval of* 1 Å *at the centre of the box, when it is in its state of least energy.*

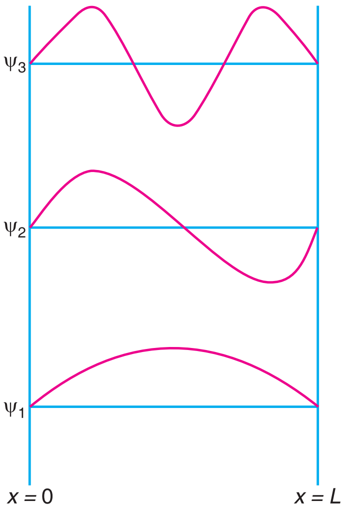


Fig. 11.15

Sol. The wave function of the particle in the ground state ($n = 1$) is

$$\psi_1 = \sqrt{\frac{2}{L}}\,\sin\frac{\pi x}{L}$$

The probability of finding the particle in unit interval at the centre of the box ($x = L/2$) is

$$P = \psi_1^2 = \left[\sqrt{\frac{2}{L}}\,\sin\frac{\pi(L/2)}{L}\right]^2 = \frac{2}{L}\sin^2\frac{\pi}{2} = \frac{2}{L}.$$

$\therefore$ The probability of finding the particle within an interval of $\Delta\,x$ at the centre of the box $= W = |\psi_1|^2\ \Delta x = \dfrac{2}{L}\Delta x$

Here, $L = 10\times10^{-10}\ m$ and $\Delta\,x = 10^{-10}\ m$.

$$\therefore \quad W = \frac{2}{10\times10^{-10}}\times10^{-10} = 0.2.$$

Example 3. *Calculate the expectation value* $<p_x>$ *of the momentum of a particle trapped in a one-dimensional box.*

Sol. The normalized wave functions of the particle are

$$\psi_n{}^* = \psi_n = \sqrt{\frac{2}{L}}\,\sin\frac{n\pi x}{L}$$

$$\frac{d\psi}{dx} = \sqrt{\frac{2}{L}}\left(\frac{n\pi}{L}\right)\cos\frac{n\pi x}{L}$$

Now,

$$<p_x> = \int_{-\infty}^{\infty} \psi^* \left(-i\frac{h}{2\pi}\frac{d}{dx}\right)\psi \, dx$$

$$= -\frac{ih}{2\pi}\frac{2}{L}\frac{n\pi}{L}\int_0^L \sin\frac{n\pi x}{L}\cos\frac{n\pi x}{L}\,dx$$

$$= 0$$

The expectation value $<p_x>$ of the particle's momentum is 0.

Example 4. *Find the expectation value $<x>$ of the position of a particle trapped in a box L wide.*

Sol.

$$< x > = \int_{-\infty}^{\infty} x\,|\psi|^2\,dx = \frac{2}{L}\int_0^L x\sin^2\frac{n\pi x}{L}\,dx$$

$$= \frac{2}{L}\left[\frac{x^2}{4} - \frac{x\sin(2n\pi x/L)}{4n\pi/L} - \frac{\cos(2n\pi x/L)}{8(n\pi/L)^2}\right]_0^L$$

$$\therefore \qquad < x > = \frac{2}{L}\left(\frac{L^2}{4}\right) = \frac{L}{2}.$$

This result means that the average position of the particle is the middle of the box in all quantum states.

11.11 Potential Step

The potential function of a potential step is defined by

$$V(x) = 0 \qquad x < 0$$

$$= V_0 \qquad x > 0 \qquad \text{...(1)}$$

Let electrons of energy E move from left to right, *i.e.*, along the positive direction of x-axis (Fig. 11.16). It is desired to find the eigenfunction solutions of the time-independent Schrodinger equation

$$\frac{d^2\psi}{dx^2} + \frac{2m}{\hbar^2}(E - V)\,\psi = 0 \qquad \text{...(2)}$$

For I region $V(x) = 0$. Therefore, the Schrodinger equation takes the form

$$\frac{d^2\psi}{dx^2} + \frac{2mE}{\hbar^2}\psi = 0 \qquad \text{...(3)}$$

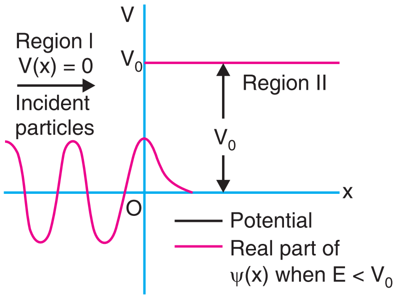


Fig. 11.16

The solution of Eq. (3) is

$$\psi_1 = Ae^{ip_1 x/\hbar} + Be^{-ip_1 x/\hbar} \qquad \text{...(4)}$$

where A and B are constants.

$$p_1 = \sqrt{(2mE)}.$$

Some particles may be reflected by the potential barrier and some transmitted. The first and second terms respectively represent the *incident* and *reflected* particles.

The Schrodinger wave equation for II region is

$$\frac{d^2\psi}{dx^2} + \frac{2m}{\hbar^2}(E - V_0)\,\psi = 0 \qquad \text{...(5)}$$

The solution of Eq. (5) is

$$\psi_2 = Ce^{ip_2x/\hbar} + De^{-ip_2x/\hbar} \qquad \text{...(6)}$$

where $p_2 = \sqrt{[2m(E - V_0)]}$; C and D are constants.

In Eq. (6), the first term represents the *transmitted wave.* The second term represents a wave coming from $+\infty$ in the negative direction. Clearly for $x > 0$ no particles can flow to the left and D must be zero. Therefore, Eq. (6) becomes

$$\psi_2 = Ce^{ip_2x/\hbar} \qquad \text{...(7)}$$

The continuity of ψ implies that $\psi_1 = \psi_2$ at $x = 0$

$$\therefore \qquad A + B = C. \qquad \text{...(8)}$$

The continuity of $\dfrac{d\psi}{dx}$ implies that $\dfrac{d\psi_1}{dx} = \dfrac{d\psi_2}{dx}$ at $x = 0$.

$$\therefore \qquad p_1(A - B) = p_2C \qquad \text{...(9)}$$

Solving (8) and (9) we get

$$B = \frac{p_1 - p_2}{p_1 + p_2}A \qquad \text{...(10)}$$

and

$$C = \frac{2p_1}{p_1 + p_2}A \qquad \text{...(11)}$$

B and C represents the amplitudes of *reflected* and *transmitted* beams respectively in terms of the amplitude of the incident wave.

The *reflectance* and the *transmittance* at the potential discontinuity may be defined as follows:

$$\text{Reflectance} \quad R = \frac{\text{magnitude of reflected current}}{\text{magnitude of incident current}}$$

$$\text{Transmittance} \quad T = \frac{\text{magnitude of transmitted current}}{\text{magnitude of incident current}}$$

Two cases may arise : (*i*) $E > V_0$ and (*ii*) $E < V_0$

Case (i): $E > V_0$. When $E > V_0$, $p_2 = \sqrt{[2m(E - V_0)]}$ is real.

We will now derive expressions for the current density in the I and II regions.

The *probability current* is defined as

$$J = \frac{\hbar}{2im}[\psi^*\nabla\psi - \psi\nabla\psi^*] \qquad \text{...(12)}$$

$$\therefore \qquad (J_x)_1 = \frac{\hbar}{2im}\left[\psi^*{}_1\frac{d\psi_1}{dx} - \psi_1\frac{d\psi_1^*}{dx}\right]$$

$$= \frac{\hbar}{2im}\left[\left\{\left(A^*e^{-ip_1x/\hbar} + B^*e^{ip_1x/\hbar}\right) \times \left(\frac{ip_1}{\hbar}\right)\left(Ae^{ip_1x/\hbar} - Be^{-ip_1x/\hbar}\right)\right\}\right.$$

$$-\left\{\left(Ae^{ip_1x/\hbar}+Be^{-ip_1x/\hbar}\right)\times\left(-\frac{ip_1}{\hbar}\right)\left(A^*e^{-ip_1x/\hbar}-B^*e^{ip_1x/\hbar}\right)\right\}\Bigg]$$

$$=\frac{p_1\,(AA^*-B^*B)}{m}=\frac{p_1}{m}\left[|A|^2-|B|^2\right] \qquad ...(13)$$

From the above relation it is evident that the current in the I region is equal to the difference between two terms. The first term which is proportional to $p_1\,|A|^2$ represents the *incident wave.* The second term which is proportional to $p_1|B|^2$ represents the *reflected wave.*

$$\left.\begin{array}{l}\text{The probability current}\\ \text{of the incident beam}\end{array}\right\}=|A|^2\,\frac{p_1}{m} \qquad ...(14)$$

$$\left.\begin{array}{l}\text{The probability current}\\ \text{of the reflected beam}\end{array}\right\}=|B|^2\,\frac{p_1}{m} \qquad ...(15)$$

The expression for the probability current in region II is

$$(J_x)_{\text{II}}=\frac{\hbar}{2im}\left[\psi_2{}^*\frac{d\psi_2}{dx}-\psi_2\frac{d\psi_2^*}{dx}\right]$$

$$=\frac{\hbar}{2im}\left[\left\{C^*e^{-ip_2x/\hbar}\left(\frac{ip_2}{\hbar}\right)Ce^{ip_2x/\hbar}\right\}\right.$$

$$\left.-\left\{Ce^{ip_2x/\hbar}\left(-\frac{ip_2}{\hbar}\right)C^*e^{-ip_2x/\hbar}\right\}\right]$$

$$=\frac{p_2}{2m}\left[CC^*+CC^*\right]=\frac{p_2}{m}\,(CC^*)$$

$$=\frac{|C|^2\,p_2}{m} \qquad ...(16)$$

Eq. (16) represents the *transmitted current.*

$$R=\frac{\text{magnitude of reflected current}}{\text{magnitude of incident current}}$$

$$=\frac{|B|^2\,(p_1/m)}{|A|^2\,(p_1/m)}$$

$$R=\frac{(p_1-p_2)^2}{(p_1+p_2)^2}\text{ from Eq. (10)} \qquad ...(17)$$

$$T=\frac{\text{magnitude of transmitted current}}{\text{magnitude of incident current}}$$

$$=\frac{|C|^2\,(p_2/m)}{|A|^2\,(p_1/m)}$$

$$=\left(\frac{2\,p_1}{p_1+p_2}\right)^2\times\frac{p_2}{p_1}\text{ from Eq. (11)}$$

$$\therefore \qquad T = \frac{4p_1p_2}{(p_1+p_2)^2} \qquad \text{...(18)}$$

Case (ii). $E < V_0$. When $E < V_0, p_2 = \sqrt{[2m(E-V_0)]}$ is imaginary.

Hence $p_2 = i\sqrt{[2m(V_0 - E)]}$ and $p_2^* = -i\sqrt{[2m(V_0-E)]} = -p_2$.

The probability current in this case is given by

$$J_x = \frac{\hbar}{2im}\left[\psi_2^* \frac{d\psi_2}{dx} - \psi_2 \frac{d\psi_2^*}{dx}\right]$$

$$= \frac{\hbar}{2im}\left[C^* e^{-ip_2^* x/\hbar}\left(\frac{ip_2}{\hbar}\right) C e^{ip_2x/\hbar} - Ce^{ip_2x/\hbar}\left(-\frac{ip_2^*}{\hbar}\right) C^* e^{-ip_2^* x/\hbar}\right]$$

Substituting $p_2{}^* = -p_2$ we get,

$$J_x = \frac{\hbar}{2im}\left[C^* e^{ip_2x/\hbar}\left(\frac{ip_2}{\hbar}\right) C e^{ip_2x/\hbar} - CC^*\left(\frac{ip_2}{\hbar}\right) e^{ip_2x/\hbar} e^{ip_2x/\hbar}\right]$$

$$= 0$$

Thus the transmitted current is zero.

$$T = \frac{\text{magnitude of transmitted current}}{\text{magnitude of incident current}} = 0$$

$$\therefore \qquad \text{T} = 0 \qquad \text{...(19)}$$

By definition, R + T = 1

$$\therefore \qquad R = 1 \qquad \text{...(20)}$$

11.12 The Barrier Penetration Problem

Consider a beam of particles of kinetic energy E incident from the left on a potential barrier of height V and width $OA = L$ (Fig. 11.17). $V > E$ and on both sides of the barrier, $V = 0$, which means that no forces act upon the particles there. This potential is described by

$V = 0$ for $x < 0$ (region I)
$V = V$ for $0 < x < L$ (region II)
$V = 0$ for $x > L$ (region III) ...(1)

Let ψ_1, ψ_2, and ψ_3 be the respective wave functions in regions I, II and III as indicated in the figure.

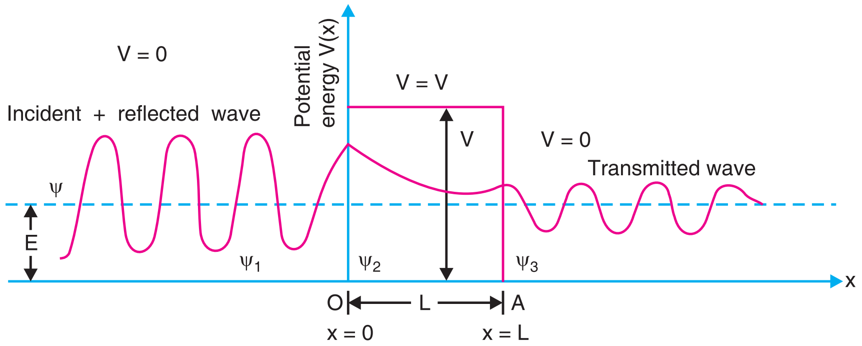


Fig. 11.17

The corresponding Schrodinger equations are

region I $\quad \dfrac{d^2\psi_1}{dx^2}+\dfrac{8\pi^2 mE}{h^2}\psi_1 = 0 \quad$ since $V = 0$

region II $\quad \dfrac{d^2\psi_2}{dx^2}+\dfrac{8\pi^2 m}{h^2}(E-V)\psi_2 = 0 \quad (\because\ V = V)$

region III $\quad \dfrac{d^2\psi_3}{dx^2}+\dfrac{8\pi^2 mE}{h^2}\psi_3 = 0 \quad (\because\ V = 0)$ ...(2)

Put $\dfrac{8\pi^2 mE}{h^2} = \alpha^2$ and $\dfrac{8\pi^2 m(V-E)}{h^2} = \beta^2$.

Then the equations become

region I $\quad \dfrac{d^2\psi_1}{dx^2}+\alpha^2\psi_1 = 0$

region II $\quad \dfrac{d^2\psi_2}{dx^2}-\beta^2\psi_2 = 0$ ...(3)

region III $\quad \dfrac{d^2\psi_3}{dx^2}+\alpha^2\psi_3 = 0$

The solutions to these equations are

region I $\quad \psi_1 = Ae^{i\alpha x}+Be^{-i\alpha x}$

region II $\quad \psi_2 = Fe^{-\beta x}+Ge^{\beta x}$ ...(4)

region III $\quad \psi_3 = Ce^{i\alpha x}+De^{-i\alpha x}$

where the constants A, B and so on are the amplitudes of the corresponding components of each wave. They may be recognized as follows:

A is the amplitude of the wave, incident on the barrier from the left,

B is the amplitude of the reflected wave in region I,

F is the amplitude of the wave, penetrating the barrier in region II,

G is the amplitude of the reflected wave (from the surface at A) in region II,

C is the amplitude of the transmitted wave, in region III, and

D is the amplitude of a (nonexistent) reflected wave, in region III.

It should be noted that we have drawn the wave function through the three regions in Fig. 11.17 so that it is continuous and singly valued everywhere along the x-axis.

Since the probability density associated with a wave function is proportional to the square of the amplitude of that function, we can define the barrier transmission coefficient as

$$T = \frac{|C|^2}{|A|^2} \qquad ...(5)$$

and a reflection coeficient for the barrier surface at $x = 0$ as

$$R = \frac{|B|^2}{|A|^2} \qquad ...(6)$$

If the barrier is high, compared to the total energy of the particle, or is thick compared to the wavelength of the wave function, then the transmission coefficient becomes

$$T \approx 16\frac{E}{V}\left(1-\frac{E}{V}\right)\exp\left[-\frac{2L}{(h/2\pi)}\sqrt{2m\,(V-E)}\right] \qquad \text{...(7)}$$

where L is the physical thickness of the barrier. The ratio $\frac{|\quad|}{|\quad|}$ is also called the 'penetrability' of the barrier. It represents the probability that a particle incident on the barrier from one side will appear on the other side. Such a probability is zero classically, but a finite quantity in quantum mechanics. We thus conclude that if a particle with energy E is incident on a thin energy barrier of height greater than E, there is a finite probability of the particle penetrating the barrier. This phenomenon is called the *tunnel effect*. This effect was used by George Gamow in 1928 to explain the process of α-decay exhibited by radioactive nuclei.

EXAMPLE. *The potential barrier problem is a good approximation to the problem of an electron trapped inside but near the surface of a metal. Calculate the probability of transmission that a* 1.0 *eV electron will penetrate a potential barrier of* 4.0 *eV when the barrier width is* 2.0 Å.

SOL. From equation (7) the transmission coefficient is

$$T \approx 16\left(\frac{1.0\,eV}{4.0\,eV}\right)\left(1-\frac{1.0\,eV}{4.0\,eV}\right)$$

$$\times\exp\left[-\frac{2\times 2\times 10^{-10}\,m}{1.05\times 10^{-34}\,Js}\sqrt{2(9.1\times 10^{-31}\,kg)(4-1)(1.6\times 10^{-19}\,J)}\right]$$

$$\approx 0.084$$

Thus, only about eight 1.0 eV electrons, out of every hundred, penetrate the barrier.

Calculation of Transmission Coefficient T

The arbitrary constants A, B, F, G and C are found from the boundary conditions at $x = 0$ and $x = L$.

The wave function and its derivative must be continuous at $x = 0$ and $x = L$.

$$\psi_1 = A\,e^{i\alpha x} + B\,e^{-i\alpha x}$$

$$\psi_2 = F\,e^{-\beta x} + G\,e^{\beta x}$$

$$\psi_3 = C\,e^{i\alpha x}$$

At $x = 0,\quad \psi_1(0) = \psi_2(0). \quad \therefore \quad A + B = F + G$ ...(1)

At $x = 0,\quad \psi_1'(0) = \psi_2'(0). \quad \therefore \quad i\alpha(A-B) = -\beta(F-G)$

or $$A - B = \frac{i\beta}{\alpha}(F-G) \qquad \text{...(2)}$$

At $x = L,\quad \psi_2(L) = \psi_3(L) \quad \therefore \quad F\,e^{-\beta L} + G\,e^{\beta L} = C\,e^{i\alpha L}$ ...(3)

At $x = L,\quad \psi_2'(L) = \psi_3'(L)$

$\therefore \quad -\beta\,[F\,e^{-\beta L} - G\,e^{\beta L}] = i\alpha\,C\,e^{i\alpha L}$ ...(4)

Solving for F and G we get

$$F = \frac{C}{2}e^{i\alpha L}\left(1-\frac{i\alpha}{\beta}\right)e^{\beta L} = \frac{C}{2\beta}(\beta - i\alpha)\,e^{i\alpha L}\,e^{\beta L}$$

$$G = \frac{C}{2}e^{i\alpha L}\left(1+\frac{i\alpha}{\beta}\right)e^{-\beta L} = \frac{C}{2\beta}(\beta + i\alpha)\,e^{i\alpha L}\,e^{-\beta L}$$

Using Eqs. (1) and (2) we get

$$A = \frac{C}{4} e^{i\alpha L}\left[-\frac{(\beta - i\alpha)^2}{i\alpha\beta} e^{\beta L} + \frac{(\beta + i\alpha)^2}{i\alpha\beta} e^{-\beta L}\right]$$

$$\frac{C}{A} = \frac{4i\alpha\beta e^{-i\alpha L}}{(\beta + i\alpha)^2 e^{-\beta L} - (\beta - i\alpha)^2 e^{\beta L}}$$

$$= \frac{4i\alpha\beta e^{-i\alpha L}}{2\left(\alpha^2 - \beta^2\right) \sin h\, \beta L + 4i\alpha\beta \cos h\, \beta L}$$

$$\therefore \quad \left|\frac{C}{A}\right|^2 = \frac{4\alpha^2\beta^2}{\left(\alpha^2 - \beta^2\right)^2 \sin h^2\, \beta L + 4\alpha^2\beta^2 \cos h^2\, \beta L}$$

$$= \frac{4\alpha^2\beta^2}{4\alpha^2\beta^2 + \left(\alpha^2 + \beta^2\right)^2 \sin h^2\, \beta L}$$

$$\therefore \quad T = \left|\frac{C}{A}\right|^2 = \left[1 + \frac{\left(\alpha^2 + \beta^2\right)^2}{4\alpha^2\beta^2} \sin h^2\, \beta L\right]^{-1}$$

$$\text{But} \quad (\alpha^2 + \beta^2)^2 = \left(\frac{8\pi^2 m}{h^2}\right)^2 V^2, \quad \alpha^2 \beta^2 = \left(\frac{8\pi^2 m}{h^2}\right)^2 E(V - E)$$

$$\therefore \quad T = \left[1 + \frac{V^2}{4E(V - E)} \sin h^2\, \beta L\right]$$

If βL is large, then $\sin h\, \beta L$ behaves as $e^{\beta L}/2$.

$$\sin h^2\, \beta L \approx \frac{1}{4} e^{2\beta L} >> 1; \quad \beta = \frac{\sqrt{2m(V - E)}}{(h/2\pi)}$$

$$\therefore \quad T = 16\frac{E}{V}\left(1 - \frac{E}{V}\right) e^{-\frac{2L}{(h/2\pi)}\sqrt{2m(V - E)}}$$

11.13 Linear Harmonic Oscillator

Consider a particle executing simple harmonic motion along the x direction and let k be the restoring force per unit displacement.

The P.E. of the particle $= \int_0^x kx\, dx = \frac{1}{2}kx^2$.

The Schrodinger equation for the harmonic oscillator is

$$\frac{d^2\psi}{dx^2} + \frac{8\pi^2 m}{h^2}\left(E - \frac{1}{2}kx^2\right)\psi = 0 \qquad \text{...(1)}$$

It is convenient to simplify Eq. (1) by introducing the dimensionless quantities

$$y = \left(\frac{1}{h/2\pi}\sqrt{km}\right)^{1/2} x = \sqrt{\frac{2\pi m\nu}{h/2\pi}}\, x \qquad \text{...(2)}$$

and $$\alpha = \frac{2E}{h/2\pi}\sqrt{\frac{m}{k}} = \frac{2E}{h\nu} \qquad \text{...(3)}$$

where ν is the classical frequency of the oscillation given by

$$\nu = \frac{1}{2\pi}\sqrt{\frac{k}{m}}\,.$$

In terms of y and α, Schrodinger's equation becomes

$$\frac{d^2\psi}{dy^2} + (\alpha - y^2)\psi = 0 \qquad \text{...(4)}$$

To solve this equation, a solution of the form below can be tried:

$$\psi = f(y)e^{-y^2/2} \qquad \text{...(5)}$$

where $f(y)$ is a function of y that remains to be found.

By inserting the ψ of Eq. (5) in Eq. (4) we obtain

$$\frac{d^2 f}{dy^2} - 2y\frac{df}{dy} + (\alpha - 1)f = 0. \qquad \text{...(6)}$$

which is the differential equation that f obeys.

Writing $(\alpha - 1) = 2n$, Eq. (6) becomes,

$$\frac{d^2 f}{dy^2} - 2y\frac{df}{dy} + 2nf = 0. \qquad \text{...(7)}$$

This is a standard mathematical equation known as *Hermite's equation.*

The solutions of Eq. (7) are called *Hermite's Polynomials*, given by

$$H_n(y) = f(y) = (-1) \exp y^2 \frac{}{dy}[\exp(-y^2)] \qquad \text{...(8)}$$

The eigen functions of harmonic oscillator, therefore, are the following:

$$\psi_n(y) = NH_n(y)\exp(-y^2/2) \qquad \text{...(9)}$$

where N is a normalisation constant.

The eigen values (permitted values of the total energy) are given by,

$$E_n = \left(n + \frac{1}{2}\right)h\nu, \quad n = 0, 1, 2, 3, \qquad \text{...(10)}$$

The energy of a harmonic oscillator is thus quantised in steps of $h\nu$. The energy levels here are evently spaced (Fig. 11.18). We note that, when $n = 0$, $E_o = \frac{1}{2}h\nu$ which is the lowest value of energy the oscillator can have. This value is called the *zero point energy*.

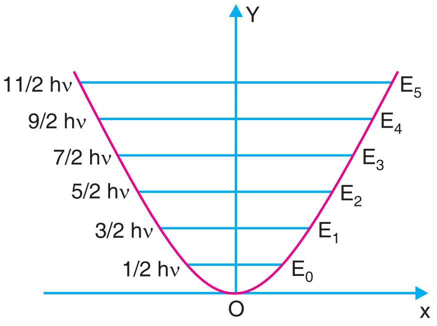


Fig. 11.18

The Harmonic Oscillator Wave Functions. Each wave function ψ_n consists of a polynomial $H_n(y)$ (called a

Hermite polynomial), the exponential factor $e^{-y^2/2}$ and a numerical coefficient which is needed for ψ_n to meet the normalisation condition

$$\int_{-\infty}^{\infty} |\psi_n|^2 \, dy = 1 \quad n = 0, 1, 2, \ldots \qquad \text{...(11)}$$

The general formula for the *n*th wave function is

$$\psi_n = \left(\frac{2m\nu}{h/2\pi}\right)^{1/4} (2^n n!)^{-1/2} H_n(y) e^{-y^2/2} \qquad \text{...(12)}$$

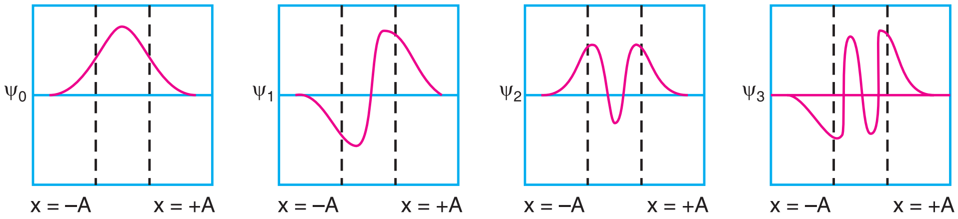


Fig. 11.19

The first four Hermite polynomials $H_n(y)$ are listed in table 11.2, and the corresponding wave functions are plotted in Fig. 11.19. The vertical lines show the limits $-A$ and $+A$ between which a classical oscillator with the same energy would vibrate.

TABLE 11.2. Some Hermite polynomials

n	$H_n(y)$	E_n
0	1	$\frac{1}{2}h\nu$
1	$2y$	$\frac{3}{2}h\nu$
2	$4y^2 - 2$	$\frac{5}{2}h\nu$
3	$8y^3 - 12y$	$\frac{7}{2}h\nu$

11.14 The Hydrogen Atom

The hydrogen atom consists of a proton around which the electron revolves. The proton is assumed at rest at the origin of a rectangular coordinate system, and the electron is oriting around the fixed proton at a radius r under the influence of the attractive coulomb field of the system. The potential energy function due to the coulomb field is $V(r) = -e^2/4\pi\varepsilon_o r$. From the viewpoint of quantum mechanics, the electron is represented by a wave system bounded by the potential well of the coulomb field. This circumstance results in a set of permitted standing wave systems, each corresponding to a particular possible value of the total energy.

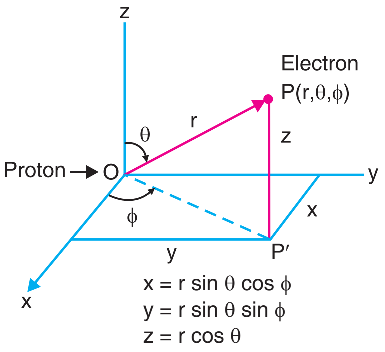


Fig. 11.20

Schrodinger's equation for the electron in three dimensions for the hydrogen atom is

$$\nabla^2 \psi + \frac{8\pi^2 m}{h^2}(E - V)\psi = 0.$$

The geometrical symmetry of the problem suggests that it is more convenient to use the spherical polar coordinate system (Fig. 11.20). The spherical polar coordinates are: (1) r, the radius vector, (2) θ, the polar angle, and (3) ϕ, the azimuth angle.

The rectangular coordinates are related to the spherical ones by

$$x = r \sin\theta \cos\phi$$

$$y = r \sin\theta \sin\phi$$

$$z = r \cos\theta \qquad \text{...(1)}$$

The spherical polar coordinates, r, θ, ϕ of the point P shown in Fig. 11.20 have the following interpretations :

r = length of radius vector from origin O to point P

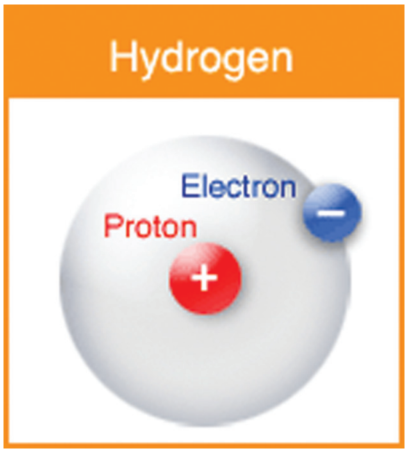
$= \sqrt{x^2 + y^2 + z^2}$

θ = angle between radius vector and + z-axis

= polar or zenith angle = $\cos^{-1} z/r$.

ϕ = angle between the projection (OP') of the radius vector OP on the xy- plane and the $+x$-axis

= azimuth angle = $\tan^{-1}(y/x)$.



Hydrogen Atom.

In spherical polar coordinates, Schrodinger's equation becomes

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial \psi}{\partial r}\right) + \frac{1}{r^2 \sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta \frac{\partial \psi}{\partial \theta}\right) + \frac{1}{r^2 \sin^2\theta}\frac{\partial^2 \psi}{\partial \phi^2} + \frac{8\pi^2 m}{h^2}(E - V)\psi = 0 \qquad \text{...(2)}$$

Substituting $V = -e^2/4\pi\epsilon_0 r$ and multiplying the entire equation by $r^2 \sin^2\theta$, we obtain

$$\sin^2\theta \frac{\partial}{\partial r}\left(r^2 \frac{\partial \psi}{\partial r}\right) + \sin\theta \frac{\partial}{\partial \theta}\left(\sin\theta \frac{\partial \psi}{\partial \theta}\right) + \frac{\partial^2 \psi}{\partial \phi^2} + \frac{8\pi^2 m r^2 \sin^2\theta}{h^2}\left(E + \frac{e^2}{4\pi\epsilon_0 r}\right)\psi = 0 \qquad \text{...(3)}$$

Equation (3) is the partial differential equation for the wave function ψ of the electron in a hydrogen atom.

Separation of Variables. The wave function is now taken to depend upon r, θ and ϕ. Here we look for solutions in which the wave function has the form of a product of three different functions.

$\psi_1(r)$, which depends upon r alone;

$\psi_2(\theta)$, which depends upon θ alone; and

$\psi_3(\phi)$, which depends upon ϕ alone.

That is, we assume that

$$\psi(r, \theta, \phi) = \psi_1(r)\, \psi_2(\theta)\, \psi_3(\phi) \qquad \text{...(4)}$$

We substitute this value of ψ in (3), divide throughout by $\psi_1 \; \psi_2 \; \psi_3$ and rearrange to get,

$$\frac{\sin^2\theta}{\psi_1}\frac{d}{dr}\left(r^2\frac{d\psi_1}{dr}\right)+\frac{\sin\theta}{\psi_2}\frac{d}{d\theta}\left(\sin\theta\frac{d\psi_2}{d\theta}\right)$$

$$+\frac{8\pi^2 mr^2\sin^2\theta}{h^2}\left(E+\frac{e^2}{4\pi\in_0 r}\right)=-\frac{1}{\psi_3}\frac{d^2\psi_3}{d\phi^2} \qquad ...(5)$$

Now the L.H.S. of this equation is a function of r and θ alone while the R.H.S. is a function of ϕ alone. Their equality implies that each side must equal a constant. It is convenient to call this constant m_l^2. Therefore, the differential equation for the function ψ_3 is

$$-\frac{1}{\psi_3}\frac{d^2\psi_3}{d\phi^2}=m_l^2.$$

This gives the *azimuthal equation* $\dfrac{d^2\psi_3}{d\phi^2}+m_l^2\,\psi_3=0$ ...(6)

When we substitute m_l^2 for the R.H.S. of Eq. (5), divide the entire equation by $\sin^2\theta$, and rearrange the various terms, we find that

$$\frac{1}{\psi_1}\frac{d}{dr}\left[r^2\frac{d\psi_1}{dr}\right]+\frac{8\pi^2 mr^2}{h^2}\left[E+\frac{e^2}{4\pi\in_0 r}\right]$$

$$=\frac{m_l^2}{\sin^2\theta}-\frac{1}{\psi_2\sin\theta}\frac{d}{d\theta}\left[\sin\theta\frac{d\psi_2}{d\theta}\right]$$

Again we have an equation in which different variables appear on each side, requiring that both sides be equal to the same constant. We shall call this constant $l\,(l+1)$. Therefore, the equations for the function ψ_2 and ψ_1 are

$$\frac{m_l^2}{\sin^2\theta}-\frac{1}{\psi_2\sin\theta}\frac{d}{d\theta}\left[\sin\theta\frac{d\psi_2}{d\theta}\right]=l\,(l+1) \qquad ...(7)$$

Eq. (7) is called the *polar equation.*

$$\frac{1}{\psi_1}\frac{d}{dr}\left(r^2\frac{d\psi_1}{dr}\right)+\frac{8\pi^2 mr^2}{h^2}\left(E+\frac{e^2}{4\pi\in_0 r}\right)=l\,(l+1) \qquad ...(8)$$

Eq. (8) is called the *radial equation.*

The Schrodinger wave equation for the hydrogen atom has now been separated into three equations [(6), (7) and (8)] each dependent on only one of the coordinates.

The Azimuthal equation. The solution of the azimuthal wave equation [Eq. (6)] is

$$\psi_3(\phi) = A\exp(im_l\phi) \qquad ...(9)$$

For this function to be single valued, its value must be the same for $\phi = \phi$ and $\phi = \phi + 2\pi$ or

$$A\exp(im_l\phi) = \text{A}\exp[im_l(\phi+2\pi)].$$

For satisfying this condition, m_l is zero or a positive or negative integer. Therefore, the only permitted values of m_l are

$$m_l = 0, \pm1, \pm2, \pm3, ... \qquad ...(10)$$

The constant m_l is known as the *magnetic quantum number* of the hydrogen atom.

The polar equation. The differential equation (7) has a rather complicated solution in terms of polynomials called the *associated Legendre polynomials* represented by $P_l m_l$ (cos θ). These polynomials exist only when the constant l is an integer equal to or greater than $|m_l|$. This requirement can be expressed as a condition on m_l in the form

$$m_l = 0, \pm 1, \pm 2, \pm 3, \ldots\ldots \pm l \qquad \ldots(11)$$

The constant l is known as the *orbital quantum number.*

The radial equation. The solutions of the radial equation (8) are given by standard mathematical functions known as the *associated Laguerre polynomials* $L_n, l\ (r)$. Equation (8) can be solved only when E is positive or has one of the negative values E_n (signifying that the electron is bound to the atom) specified by

$$E_n = -\frac{me^4}{8\epsilon_0^2 h^2}\left(\frac{1}{n^2}\right) = (-13.6)\frac{1}{n^2} eV \qquad \ldots(12)$$

This is in complete agreement with the predictions of the Bohr theory. Here n is again an integer whose value must be greater than or equal to $(l + 1)$ *i.e.*, $n \geq (l + 1)$. We have now arrived at a system of three quantum numbers for the simple model of the hydrogen atom, which are related in the following way.

total quantum number	$n = 1, 2, 3, \ldots\ldots\ldots$
orbital quantum number	$l = 0, 1, 2, \ldots\ldots (n-1)$
magnetic quantum number	$m_l = 0, \pm 1, \pm 2, \ldots\ldots\ldots \pm l$

11.15 The Rigid Rotator

Consider two point masses M_1 and M_2 joined together by a rigid, massless rod [Fig. 11.21]. The two masses are rotating like a dumbbell about an axis through their centre of mass C normal to the line joining the two masses. Using the relations $M_1 r_1 = M_2 r_2$, and $r_1 + r_2 = r$, we can express the moment of inertia I of the system about the axis of rotation in terms of r and get

$$I = M_1 r_1^2 + M_2 r_2^2 = \frac{M_1 M_2}{M_1 + M_2} r^2 = \mu r^2$$

where $\mu = M_1 M_2 / M_1 + M_2$ is the *reduced mass* of the system. We can thus replace the rotating dumbbell by a single point mass μ at a distance r from the axis of rotation. We can take the P.E. of the rotating dumbbell as zero, since the two point masses are rigidly joined together, with the distance between them remaining unaffected due to the rotation. The angular momentum and energy of this *simple rigid rotator* are given respectively, by $p = I\omega$ and $E = \frac{1}{2} I\omega^2$, where ω is the uniform angular speed of rotation. According to quantum theory, the angular momentum should be an integral multiple of $h/2\,\pi$. Hence we get

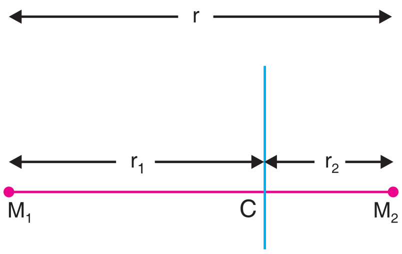


Fig. 11.21

$$p = I\omega = kh/2\pi, \quad \text{and} \quad E = E_k = \frac{1}{2} I\omega^2 = \frac{k^2 h^2}{8\pi^2 I}$$

where k is an integer.

To derive the permitted energy values of the rigid rotator on the wave mechanical theory. The general Schrodinger equation in spherical polar coordinates is given by

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial \psi}{\partial r}\right) + \frac{1}{r^2}\frac{1}{\sin\theta}\frac{\partial}{\partial \theta}\left(\sin\theta \frac{\partial \psi}{\partial \theta}\right)$$

$$+ \frac{1}{r^2 \sin^2\theta}\frac{\partial^2 \psi}{\partial \phi^2} + \frac{8\pi^2 \mu}{h^2}(E - V)\,\psi = 0 \qquad \ldots(1)$$

As r is constant ($r = 1$), the first factor $\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right) = 0$, $V = 0$ and μ may be replaced by I.

Schrodinger equation of rigid rotator reduces to

$$\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2\psi}{\partial\phi^2}+\frac{8\pi^2 I}{h^2}E\psi = 0 \qquad ...(2)$$

The solution of equation (2) can be obtained by the method of separation of variables. Assume

$$\psi(\theta,\phi) = \psi_1(\theta)\,\psi_2(\phi) \qquad ...(3)$$

Substituting (3) in equation (2), and dividing both sides by $\psi_1\,\psi_2/\sin^2\theta$, we have

$$\frac{\sin\theta}{\psi_1}\frac{d}{d\theta}\left(\sin\theta\frac{d\psi_1}{d\theta}\right)+\frac{8\pi^2 IE}{h^2}\sin^2\theta = -\frac{1}{\psi_2}\frac{d^2\psi_2}{d\phi^2} \qquad ...(4)$$

Define a quantity m^2 by the following equation

$$\frac{1}{\psi_2}\frac{d^2\psi_2}{d\phi^2} = -m^2 \textit{ i.e., } \frac{d^2\psi_2}{d\phi^2}+m^2\psi_2 = 0 \qquad ...(5)$$

$$\therefore \qquad \frac{\sin\theta}{\psi_1}\frac{d}{d\theta}\left(\sin\theta\frac{d\psi_1}{d\theta}\right)+\frac{8\pi^2 IE}{h^2}\sin^2\theta = m^2$$

or

$$\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{d\psi_1}{d\theta}\right)+\left(\frac{8\pi^2 IE}{h^2}-\frac{m^2}{\sin^2\theta}\right)\psi_1 = 0 \qquad ...(6)$$

The solution of equation (5) is

$\psi_2 = \exp(\pm im\phi)$ where m is any positive integer.

Putting $\frac{8\pi^2 IE}{h^2} = C$, equation (6) becomes

$$\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{d\psi_1}{d\theta}\right)+\left(C-\frac{m^2}{\sin^2\theta}\right)\psi_1 = 0 \qquad ...(7)$$

This equation can be solved by substituting $\cos\theta = x$

Equation (7) is then transformed to

$$\frac{d}{dx}\left[(1-x^2)\frac{d\psi_1}{dx}\right]+\left[C-\frac{m^2}{1-x^2}\right]\psi_1 = 0$$

Let us try the solution $\psi_1 = (1-x^2)^{m/2}P(x)$ ... (8)

where $P(x)$ is a function of x to be determined. We choose this type of solution, since it enables us to put the above differential equation in a standard form, namely

$$(1-x^2)\frac{d^2P}{dx^2}-2x(m+1)\frac{dP}{dx}+(C-m-m^2)P = 0.$$

We attempt a solution of this equation by expressing P as a power series in x, namely

$$P = \sum_{P=-\infty}^{+\infty} a_p x^p$$

Substituting this value of P in the differential equation, we get

$$\Sigma p\,(p-1)\,a_p x^{p-2}-\Sigma\,[p\,(p-1)+2\,(m+1)\,p-(C-m-m^2)]\,a_p x^p = 0.$$

We find from the recurrence relation between the coefficients in the above equation that solutions in closed form are obtained when

$C = (p + m)\,(p + m + 1) = k\,(k + 1),$

say, where we have put $k = p + m$.

Substituting $C = k\,(k+1)$ in $C = \dfrac{8\pi^2 IE}{h^2}$ we get finally

$$E = k\,(k+1)\,\frac{h^2}{8\pi^2 I} \qquad \text{...(9)}$$

k is called the *rotational quantum number*.

Eq. (9) gives allowed values for the energy (*i.e.*, eigen values) of a rigid rotator with free axis.

EXERCISE

1. Discuss briefly the wave nature of matter and obtain an expression of de Broglie wavelength for matter waves.

2. Describe Davisson and Germer experiments for the study of electron diffraction. What are the results of the experiment?

3. Describe briefly the experiment of G.P. Thomson on the diffraction of electrons. Explain briefly the results obtained.

4. What is meant by matter waves? Give experimental evidence in support of the concept of these waves.

5. Give an account of Heisenberg's uncertainty Principle. Outline an idealised experiment to bring out its significance.

6. Explain the principles on which the electron microscope works. Compare this instrument with an optical microscope.

7. What are the postulates of wave mechanics ? Derive Schrodinger's equation. What is the significance of the wave function?

8. Calculate the values of the energy of a particle in a one dimensional box. Indicate graphically the first three wave functions for such a particle.

9. Explain the requirements that are imposed on a physically acceptable wave function. What do you understand by the terms "*Eigen value*" and "*Eigen function*"?

10. Explain what you understand by the terms potential well and potential barrier. How does a particle with energy lower than the barrier height, tunnel through it? Give one example.

11. Establish Schrodinger's equation for a linear harmonic oscillator and solve it to obtain its eigen values and eigen functions. In what way is the quantum mechanical description of a simple harmonic oscillator different from classical description?

12. Write the Schrodinger wave equation of the hydrogen atom in polar coordinates. Explain the origin and significance of the quantum numbers n, l and m_l.

13. Solve the Schrodinger's equation for the case of hydrogen atom. Neglect electron spin and assume the nucleus to be stationary.

14. Formulate Schrodinger's equation for a rigid rotator. Find its eigen values and eigen functions.

15. Determine the wavelength associated with an electron having *K.E.* equal to 1 *MeV*.

[**Hint.** Since the K.E. of the electron (1 MeV) is comparable with its rest mass energy (0.512 MeV), we must use relativistic equation. Thus $mc^2 = m_0c^2 + \text{K.E.} = 1.512\ \text{MeV} = 1.512 \times (1.6 \times 10^{-13})$ J or $m = 3.18 \times 10^{-30}$ kg. Now, $m = m_0/\sqrt{1-v^2/c^2}$

Here, $m_o = 9.1 \times 10^{-31}$kg. $\therefore\ v = 2.87 \times 10^8\ \text{ms}^{-1}$.

$$\therefore \qquad \lambda = \frac{h}{mv} = \frac{6.6 \times 10^{-34}}{(3.18 \times 10^{-30})\,(2.87 \times 10^8)}\ m = 0.00726\ \text{Å}].$$

16. A proton and an α-particle have the same K.E. (*a*) How do their speeds compare? (*b*) How do their momenta compare? (*c*) How do their de Broglie wavelengths compare? Given $m_\alpha = 4m_p$. [**Ans.** (*a*) $v_p = 2v_\alpha$ (*b*) $P_p = \frac{1}{2}\, p_\alpha$ (*c*) $\lambda_p = 2\lambda_\alpha$]

17. Calculate the wavelength associated with electrons whose speed is 0.01 the speed of light. ($h = 6.62 \times 10^{-34}$ joule-sec: $m = 9.11 \times 10^{-31}$ kg). [**Ans.** $2.4 \times 10^{-10}\ m$]

18. Calculate the de Broglie wavelength of an α particle accelerated by a P.D. of 25000 volts. [**Ans.** $6.390 \times 10^{-14}\ m$]

19. Calculate the glancing angle at which electrons of energy 100 eV must be incident on the lattice planes of a metal crystal in order to give a strong Bragg reflection in the first order. (Lattice spacing = 2.52 Å). [**Ans.** 14°8’]

20. The equivalent wavelength of a moving electron is $0.24 \times 10^{-10}\ m$. What voltage applied between two grids will bring it to rest ?

[**Hint.**
$$mv = \frac{h}{\lambda} \quad \text{or} \quad v = \frac{h}{\lambda m} = \frac{6.626 \times 10^{-34}}{(0.24 \times 10^{-10})\,(9.109 \times 10^{-31})}$$
$$= 3.032 \times 10^{7}\ ms^{-1}$$
$$\therefore \quad \frac{1}{2} mv^2 = \frac{1}{2}(9.109 \times 10^{-31})\,(3.032 \times 10^{7})^2 = 4.186 \times 10^{-16}\ J$$
$$= 2616\ eV$$
∴ The voltage required is 2616 *V*.]

21. A beam of mono-energetic neutrons corresponding to 27°*C* is allowed to fall on a crystal. A first order reflection is observed at a glancing angle 30°; calculate the interplanar spacing of the crystal. Given: $h = 6.62 \times 10^{-34}$ js; $m_n = 1.67 \times 10^{-27}$ kg. Boltzmann’s constant, $k = 1.38 \times 10^{-23}$ joule/degree.

[**Hint.** $2d \sin\theta = n\lambda$; Here, $\theta = 30°$ and $n = 1$. $\therefore\ \lambda = d$.

Energy of neutron = $E = kT$

We have $d = \lambda = \dfrac{h}{p} = \dfrac{h}{(2mE)^{1/2}}$
$$= \frac{6.62 \times 10^{-34}}{[2\,(1.67 \times 10^{-27})\,1.38 \times 10^{-23} \times 300]^{1/2}} = 1.82 \times 10^{-10}\,m]$$

22. What bearing would you think the uncertainty principle has on the existence of the zero-point energy of a harmonic oscillator?

[**Ans.** The oscillator cannot have zero energy because this would mean it is at rest in a definite position, whereas according to the uncertainty principle a definite position corresponds to an infinite momentum (and hence energy) uncertainty].

23. A particle is diffracted at a certain angle and the uncertainty in the measurement of the angle is one second of arc. Estimate the error in the measurement of its angular momentum.

[**Hint.** $\Delta\phi$ = one second = $\pi/60 \times 60 \times 180$ radian.
$$\therefore \quad \Delta L = \frac{(h/2\pi)}{\Delta\phi} = 2.168 \times 10^{-29}\ Js].$$

24. Does the concept of Bohr orbit violate the Heisenberg’s uncertainty principle?

[**Ans.** According to Bohr, an electron in an atom revolves in one of the quantised orbits. The electron energy $E_n = -e^2/(8\pi\epsilon_o\, r_n)$. An atomic electron can have only these energies and no others. *i.e.*, $\Delta E = 0$. According to uncertainty relation, $\Delta E\, \Delta t \geq \dfrac{h}{2\pi}$. Therefore, Δt must be infinite which means that all energy states of the atom must have an infinite lifetime. But we know that the excited states of the atom from which radiation occurs do not last indefinitely and the life-time is of the order of 10^{-8} sec. Thus the concept of Bohr orbit violates the uncertainty principle].

25. A bullet of mass 25 grams is moving with a speed of 400 m/s. The speed is measured accurate up to 0.02%. Calculate the certainty with which the position of the bullet can be located. $h = 6.6 \times 10^{-34}$ joule second [**Ans.** $\Delta x \geq 5.25 \times 10^{-32} m$]

26. The position and momentum of a 1 KeV electron are simultaneously determined. If its position is located to within 1Å, what is the percentage of uncertainty in its momentum?

[**Hint.** $\Delta p = \dfrac{h}{2\pi \Delta x} = \dfrac{6.626 \times 10^{-34}}{2\pi \times 10^{-10}}$

$$= 1.054 \times 10^{-24}\ kg\ ms^{-1}$$

$$p = mv = \sqrt{2meV} = \sqrt{2(9.109 \times 10^{-31})(1.602 \times 10^{-19})\,1000}$$

$$\therefore \quad p = 1.708 \times 10^{-23}\ \text{kg ms}^{-1}$$

$$\therefore \quad \frac{\Delta p}{p} \times 100 = \frac{1.054 \times 10^{-24}}{1.708 \times 10^{-23}} \times 100 = 6.173\%]$$

27. Compare the uncertainties in the velocities of an electron and a proton confined in a 10 Å box.

[**Ans.** 1.16×10^5 ms^{-1}; 63 ms^{-1}]

28. Applying uncertainty principle, show that the presence of protons in a nucleus is entirely plausible.

[**Hint.** $\Delta p \geq 1.1 \times 10^{-20}$ kg ms^{-1} $\therefore$ The proton momentum must be atleast as large as the minimum value of Δp. For a proton with a momentum of 1.1×10^{-20} kg ms^{-1}, $T << m_0 c^2$ and its K.E. can be calculated classically. We have $T = p^2/2m = 3.6 \times 10^{-14}$ J = 0.23 MeV. The presence of protons with such kinetic energies in a nucleus is entirely plausible]

29. A proton is confined to a nucleus of radius 5×10^{-15} m. Calculate the minimum uncertainty in its momentum. Also calculate the minimum K.E. the proton should have. The proton mass is 1.67×10^{-27} kg.

[**Hint.** The diameter of the nucleus = $\Delta x = 10^{-14}$ m

$$\therefore \quad \Delta p = \frac{h/2\pi}{\Delta x} = 1.05 \times 10^{-20}\ kg\ ms^{-1}$$

Then $\quad p = 1.05 \times 10^{-20}$ kg ms^{-1}.

The minimum K.E. of the proton $= \dfrac{p^2}{2m_0} = 0.2$ MeV]

30. An electron is confined to move between two rigid walls separated by 10^{-9} m. Find the de Broglie wavelengths representing the first three allowed energy states of the electron and the corresponding energies.

[**Hint.** $L = n\lambda/2$ since the distance (L) between the walls must be an integral multiple of $\lambda/2$.

$$\therefore \quad \lambda = \frac{2L}{n} = \frac{2 \times 10\text{Å}}{n} \quad \text{where } n = 1, 2, 3, ...]$$

[**Ans.** (i) 20 Å, 10Å, 6.7Å (ii) 0.38 eV, 1.52 eV, 3.42 eV]

31. *Obtain Schrödinger's wave equation in momentum representation.*

Ans. In three dimensions, the time-dependent form of Schrodinger's equation is

$$\frac{-\hbar^2}{2m}\nabla^2\psi + V\psi = i\hbar\frac{\partial\psi}{\partial t} \qquad ...(1)$$

The wave function for any plane wave is given by

$$\psi = exp.\left[\frac{i}{\hbar}(\vec{p}\cdot\vec{r} - Et)\right]$$

Any superposition of such waves forms a wave packet.

The wave equation of the wave packet is

$$\psi = \frac{1}{\left(\sqrt{2\pi\hbar}\right)^3}\int a(p)exp.\left[\frac{i}{\hbar}(\vec{p}\cdot\vec{r} - Et)\right]dp \qquad ...(2)$$

Differentiating Eq. (2) with respect to t, we get

$$\frac{\partial \psi}{\partial t} = \frac{1}{\left(\sqrt{2\pi\hbar}\right)^3}\int\left(\frac{-iE}{\hbar}\right)a(p)\exp.\left[\frac{i}{\hbar}(\vec{p}\cdot\vec{r} - Et\right]dp \qquad ...(3)$$

Further $\vec{p}\cdot\vec{r} = p_x x + p_y y + p_z z$

$$\therefore \qquad \frac{\partial \psi}{\partial x} = \frac{1}{\left(\sqrt{2\pi\hbar}\right)^3}\int\frac{ip_x}{\hbar}a(p)\exp.\left[\frac{i}{\hbar}(\vec{p}\cdot\vec{r} - Et\right]dp$$

and

$$\frac{\partial^2 \psi}{\partial x^2} = \frac{1}{\left(\sqrt{2\pi\hbar}\right)^3}\int\left(\frac{-p_x^2}{\hbar^2}\right)a(p)\exp.\left[\frac{i}{\hbar}(\vec{p}\cdot\vec{r} - Et\right]dp$$

$$\frac{\partial^2 \psi}{\partial y^2} = \frac{1}{\left(\sqrt{2\pi\hbar}\right)^3}\int\left(\frac{-p_y^2}{\hbar^2}\right)a(p)\exp.\left[\frac{i}{\hbar}(\vec{p}\cdot\vec{r} - Et)\right]dp$$

and

$$\frac{\partial^2 \psi}{\partial z^2} = \frac{1}{\left(\sqrt{2\pi\hbar}\right)^3}\int\left(\frac{-p_z^2}{\hbar^2}\right)a(p)\exp.\left[\frac{i}{\hbar}(\vec{p}\cdot\vec{r} - Et\right]dp$$

$$\therefore \qquad \nabla^2\psi = \frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2}$$

$$= \frac{1}{\left(\sqrt{2\pi\hbar}\right)^3}\int\left(-\frac{p_x^2 + p_y^2 + p_z^2}{\hbar^2}\right)a(p)\exp.\left[\frac{i}{\hbar}(\vec{p}\cdot\vec{r} - Et)\right]dp$$

$$= \frac{1}{\left(\sqrt{2\pi\hbar}\right)^3}\int\left(-\frac{p^2}{\hbar^2}\right)a(p)\exp.\left[\frac{i}{\hbar}(\vec{p}\cdot\vec{r} - Et)\right]dp \qquad ...(4)$$

Substituting the value of $\nabla^2\psi$ from Eq. (4) and $\frac{\partial \psi}{\partial t}$ from Eq. (3) in Eq. (1), we get

$$-\frac{\hbar^2}{2m}\left(-\frac{p^2}{\hbar^2}\right)\frac{1}{\left(\sqrt{2\pi\hbar}\right)^3}\int a(p)\exp.\left[\frac{i}{\hbar}(\vec{p}\cdot\vec{r} - Et\right]dp + V\psi$$

$$= \frac{1}{\left(\sqrt{2\pi\hbar}\right)^3}i\hbar\left(\frac{-iE}{\hbar}\right)\int a(p)\exp.\left[\frac{i}{\hbar}(\vec{p}\cdot\vec{r} - Et\right]dp$$

or

$$\frac{1}{\left(\sqrt{2\pi\hbar}\right)^3}\left(\frac{p^2}{2m} - E\right)\int a(p)\exp.\left[\frac{i}{\hbar}(\vec{p}\cdot\vec{r} - Et)\right]dp + V\psi = 0 \qquad ...(5)$$

This is Schrödinger's three dimensional wave equation in momentum representation.

One dimensional wave equation. For a one dimensional case, Schrödinger's wave equation is

$$\frac{-\hbar}{2m}\frac{\partial^2 \psi}{\partial x^2} + V\psi = i\hbar\frac{\partial \psi}{\partial t} \qquad ...(6)$$

The equation for a wave packet in one dimension is

$$\psi = \frac{1}{\sqrt{2\pi\hbar}}\int a(p)\exp.\left[\frac{i}{\hbar}(p_x x - Et)\right]dp$$

$$\therefore \qquad \frac{\partial \psi}{\partial t} = \frac{1}{\sqrt{2\pi\hbar}}\int\left(\frac{-iE}{\hbar}\right)a(p)\exp.\left[\frac{i}{\hbar}(p_x x - Et)\right]dp \qquad ...(7)$$

and

$$\frac{\partial \psi}{\partial x} = \frac{1}{\sqrt{2\pi\hbar}}\int\left(\frac{i}{\hbar}p_x\right)a(p)\exp.\left[\frac{i}{\hbar}(p_x x - Et)\right]dp$$

$$\frac{\partial^2\psi}{\partial x^2} = \frac{1}{(\sqrt{2\pi\hbar})}\int -\left(\frac{p_x}{\hbar}\right)^2 a(p)\exp.\left[\frac{i}{\hbar}(p_x x - Et\right]dp \qquad \text{...(8)}$$

Substituting the value of $\frac{\partial^2\psi}{\partial x^2}$ from Eq, (8), and $\frac{\partial\psi}{\partial t}$ from Eq. (7) in Eq. (6), we get

$$\frac{1}{\sqrt{2\pi\hbar}}\left(\frac{p_x^2}{2m} - E\right)\int a(p)\exp.\left[\frac{i}{\hbar}(p_x x - Et\right]dp + V\psi = 0 \qquad \text{...(9)}$$

This is Schrödinger's one dimensional wave equation in momentum representation.

32. *Verify the orthogonal property of wave functions of a particle in a one dimensional box.*

Sol. The wave functions of a particle in one dimensional box are

$$\psi_n(x) = \sqrt{\frac{2}{L}}\sin\frac{n\pi x}{L}$$

We want to verify

$$\int_0^L \psi_n^*(x)\psi_m(x)dx = 0 \text{ for } n \neq m.$$

$$\text{L.H.S} = \int_0^L \sqrt{\frac{2}{L}}\sin\frac{n\pi x}{L}\sqrt{\frac{2}{L}}\sin\frac{m\pi x}{L}dx = \frac{1}{L}\int_0^L\left(2\sin\frac{n\pi x}{L}\sin\frac{m\pi x}{L}\right)dx$$

$$= \frac{1}{L}\int_0^L\left[\cos(n-m)\frac{\pi x}{L} - \cos(n+m)\frac{\pi x}{L}\right]dx$$

∴ L.H.S = 0. Hence verified.

33. *Using the uncertainty principle, estimate the ground state energy of the harmonic oscillator.*

Ans. Assume that the particle position is uncertain by Δx. Then, from the uncertainty principle, the uncertainty in particle momentum is $\Delta p = \hbar/(2\Delta x)$.

Total energy of the linear harmonic oscillator is

$$E = \frac{1}{2}k(\Delta x)^2 + \frac{(\Delta p)^2}{2m} = \frac{1}{2}k(\Delta x)^2 + \frac{\hbar^2}{8m(\Delta x)^2} \qquad \text{...(1)}$$

The value of E will be minimum when $\frac{dE}{d(\Delta x)} = 0$

$$k\Delta x - \frac{\hbar^2}{4m(\Delta x)^3} = 0 \text{ or } \Delta x = \left(\frac{\hbar^2}{4mk}\right)^{1/4}.$$

Substituting the value of Δx in Eq. (1), we get

$$E_{\min} = \frac{1}{2}k\left(\frac{\hbar^2}{4mk}\right)^{1/2} + \frac{\hbar^2}{8m}\left(\frac{4mk}{\hbar^2}\right)^{1/2} = \frac{\hbar}{2}\left(\frac{k}{m}\right)^{1/2} = \frac{\hbar\omega}{2}.$$

34. *Solve linear harmonic oscillator problem using operator methods. Obtain the energy eigenvalues and eigenfunctions.*

Ans. **(a) The Ladder (or Raising and Lowering) Operators:**

The Hamiltonian of a linear harmonic oscillator is

$$H = \frac{1}{2m}p^2 + \frac{1}{2}m\omega_c{}^2x^2 \qquad \text{...(1)}$$

ω_c is the angular frequency of the harmonic oscillator.

x and p are the coordinate and momentum operators. The quantum condition is

$$[x, p] = xp - px = i\hbar \qquad \text{...(2)}$$

Let us now introduce a pair of operators,

$$a = \left(\frac{m\omega_c}{2\hbar}\right)^{1/2} x + i\left(\frac{1}{2m\hbar\omega_c}\right)^{1/2} p \qquad \text{...(3)}$$

$$a^\dagger = \left(\frac{m\omega_c}{2\hbar}\right)^{1/2} x - i\left(\frac{1}{2m\hbar\omega_c}\right)^{1/2} p \qquad \text{...(4)}$$

$$aa^\dagger = \frac{H}{\hbar\omega_c} + \frac{1}{2} \text{ and } a^\dagger a = \frac{H}{\hbar\omega_c} - \frac{1}{2} \qquad \text{...(5)}$$

a and $a^\dagger$ satisfy the commutation relation

$$[a, a^\dagger] = aa^\dagger - a^\dagger a = 1 \qquad \text{...(6)}$$

$$H = \left(a^\dagger a + \frac{1}{2}\right)\hbar\omega_c \qquad \text{...(7)}$$

Since the relation between H and $a^\dagger a$ involves only ordinary numbers (no operators), their eigenvalues will also bear the same relation, and their eigenfunctions coincide. We shall now see that the eigenvalues of $a^\dagger a$ are integers $n (n = 0, 1, 2, ...)$. So the eigenvalues of the Hamiltonian are $\left(a^\dagger a + \frac{1}{2}\right)\hbar\omega_c$.

$$[a^\dagger a, a] = (a^\dagger a)a - a(a^\dagger a) = (a^\dagger a - aa^\dagger)a = -a \qquad \text{...(8)}$$

Similarly $$[a^\dagger a, a^\dagger] = a^\dagger \qquad \text{...(9)}$$

Let u be an eigenfunction of $(a^\dagger a)$, belonging to the eigenvalue λ, i.e.,

$$(a^\dagger a)u = \lambda u \qquad \text{...(10)}$$

Using Eqs. (6) and (10), we have

$$(a^\dagger a)au = (aa^\dagger - 1)au = a(a^\dagger a - 1)u = a(\lambda - 1)u = (\lambda - 1)au \qquad \text{...(11)}$$

Similarly, $$(a^\dagger a)a^\dagger u = (\lambda + 1)a^\dagger u \qquad \text{...(12)}$$

Eqs. (11) and (12) state that au and $a^\dagger u$ are also eigenfunctions of $(a^\dagger a)$, belonging to the eigenvalues $(\lambda - 1)$ and $(\lambda + 1)$ respectively.

Starting with u, we can now construct wave functions $a^\dagger u$, $(a^\dagger)^2 u$, ... belonging to eigenvalues $(\lambda + 1)$. $(\lambda + 2)$, ... and also au, $a^2 u$, ... belonging to $(\lambda - 1)$, $(\lambda - 2)$, ... The eigenvalues thus form a 'ladder' with unit spacing between steps.

$a^\dagger$ has the effect of raising the eigenvalue by one step.

a has the effect of lowering the eigenvalue by one step.

For this reason, $a^\dagger$ and a are called ladder operators, or *raising* and *lowering operators*.

(*b*) *The Eigenvalue Spectrum***:** $a^\dagger a$ is a positive operator. Therefore, it can have no negative eigenvalues. So the sequence of eigenvalues $(\lambda - 1)$, $(\lambda - 2)$... must terminate before it reaches negative values, and correspondingly, the sequence au, $a^2 u$, ... must terminate. Let us call the last eigenfunction in the sequence as u_0. Then, by definition, there does not exist a state au_0. In other words,

$$au_0 = 0 \qquad \text{...(13)}$$

This implies obviously that

$$a^\dagger a\, u_0 = 0 \qquad \text{...(14)}$$

i.e., u_0 is an eigenstate of $(a^\dagger a)$, belonging to the eigenvalue 0. The other eigenstates can now be constructed by repeated application of $a^\dagger$ to u_0. Since the effect of $a^\dagger$ applied once is to increase the eigenvalue by unity, we have that the eigenstates u_0, $(a^\dagger)u_0$, $(a^\dagger)^2 u_0$ ... belong to the eigenvalues, 0, 1, 2, ... of $(a^\dagger a)$.

Thus the eigenvalue spectrum of $(a^\dagger a)$ consists of the set of nonnegative integers n.

Hence $a^\dagger a$ is called the *number operator*.

The eigenvalues of H are $\left(n + \frac{1}{2}\right)\hbar\omega_c \quad n = 0, 1, 2, ...$

(*c*) *The Energy Eigenfunctions*: The eigenstates which we have constructed above are mutually orthogonal. The norm of $(a^\dagger)^n u_0$ is $n!$

$$\left((a^\dagger)^m\, u_0, (a^\dagger)^n\, u_0\right) = n\,!\,\delta_{mn} \qquad \text{...(15)}$$

Therefore, the normalized eigenfunctions are given by

$$u_n = \frac{(a^\dagger)^n}{\sqrt{n!}} u_0, \;\; n = 0, 1, 2, \ldots \qquad \text{...(16)}$$

By construction, they satisfy the relation

$$(a^\dagger a) u_n = n u_n \qquad \text{...(17)}$$

Observe that Eq. (16) can be rewritten as

$$u_n = \frac{n^\dagger}{\sqrt{n}} \frac{(a^\dagger)^{n-1}}{\sqrt{(n-1)!}} u_0 = \frac{a^\dagger}{\sqrt{n}} u_{n-1}$$

$$\therefore \qquad a^\dagger u_{n-1} = \sqrt{n}\, u_n \qquad \text{...(18)}$$

Further, by operating on this equation with a, we get

$$\sqrt{n}\, a u_n = a a^\dagger u_{n-1} = (a^\dagger a + 1) u_{n-1} = n u_{n-1}$$

$$\therefore \qquad a u_n = \sqrt{n}\, u_{n-1} \qquad \text{...(19)}$$

From Eqs. (18) and (19), we get

$$(u_m, a^\dagger u_n) = \sqrt{n+1}\, \delta_{m,\, n+1} \;\; ; (u_m, a u_n) = \sqrt{n}\, \delta_{m,\, n-1} \qquad \text{...(20)}$$

Eqs. (18), (19) and (20) enable us to calculate the effect of any function of x and p on the energy eigenfunctions by expressing them first in terms of a and $a^\dagger$ through Eqs. (3) and (4).

Determination of the actual form of the wave functions (coordinate-representation)

Consider Eq. (13). On substituting Eq. (3) in it and writing $p = -i\hbar\, d/dx$, we get the differential equation

$$\left(\frac{d}{dx} + \frac{m\omega_c}{\hbar} x\right) u_0 = 0 \qquad \text{...(21)}$$

The solution of this equation, with proper choice of normalization constant, is

$$u_0(x) = \left(\alpha^2/\pi\right)^{1/4} \cdot e^{-\frac{1}{2}\alpha^2 x^2}, \text{ with } \alpha = (m\omega_c/\hbar)^{1/2} \qquad \text{...(22)}$$

All other eigenfunctions are obtained from this by using Eq. (16).

Observing that $a^\dagger$ can be written in the form

$$a^\dagger = -\frac{1}{\sqrt{2}}\left(\frac{d}{d\rho} - \rho\right) = -\frac{1}{\sqrt{2}} e^{\frac{1}{2}\rho^2} \frac{d}{d\rho} e^{-\frac{1}{2}\rho^2}, \text{ with } \rho = \alpha x,$$

and using the fact that the Hermite polynomials can be expressed as

$$H_n(\rho) = (-)^n e^{\rho^2} \frac{d^n}{d\rho^n} e^{-\rho^2} = (-)^n e^{\frac{1}{2}\rho^2} \left(e^{\frac{1}{2}\rho^2} \frac{d}{d\rho} e^{-\frac{1}{2}\rho^2}\right)^n e^{-\frac{1}{2}\rho^2} \qquad \text{...(23)}$$

We get

$$u_n(x) = N_n e^{\frac{1}{2}\rho 2} H_n(\rho) = N_n e^{-\frac{1}{2}\alpha^2 x^2} \cdot H_n(\alpha x), n = 0, 1, 2, \ldots, \qquad \text{...(24)}$$

The normalization constant $N_n = \left(\dfrac{\alpha}{\sqrt{\pi}\, 2^n n!}\right)^{1/2}$ ...(25)

35. *Solve the polar equation for hydrogen atom.* [*H.P.U. 1999*]

Sol. The polar equation is

$$\frac{m_l^2}{\sin^2\theta} - \frac{1}{\psi_2 \sin\theta} \frac{d}{d\theta}\left(\sin\theta \frac{d\psi_2}{d\theta}\right) = l(l+1) \qquad \text{...(1)}$$

$$\text{or} \quad \frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{d\psi_2}{d\theta}\right)+\left[l(l+1)-\frac{m_l^2}{\sin^2\theta}\right]\psi_2 = 0 \qquad \text{...(2)}$$

Put $x = \cos\theta \;\therefore\; \frac{dx}{d\theta} = -\sin\theta\,;\; \frac{d\psi_2}{d\theta} = \frac{d\psi_2}{dx}\frac{dx}{d\theta} = -\sin\theta\frac{d\psi_2}{dx}$

$$\therefore \qquad \frac{\quad}{\quad} = -\sin\theta\frac{d}{dx}$$

Putting $\frac{d}{d\theta}$ and $\frac{d\psi_2}{d\theta}$ in Eq. (2), we get

$$\frac{1}{\sin\theta}\times -\sin\theta\frac{d}{dx}\left(-\sin^2\theta\frac{d\psi_2}{dx}\right)+\left[l(l+1)-\frac{m_l^2}{\sin^2\theta}\right]\psi_2 = 0 \qquad \text{...(3)}$$

$$\text{or} \quad \frac{d}{dx}\left[(1-x^2)\frac{d\psi_2}{dx}\right]+\left[l(l+1)-\frac{m_l^2}{(1-x^2)}\right]\psi_2 = 0 \qquad \text{...(4)}$$

The differential equation (4) has a solution in terms of polynomials called the *associated Legendre polynomials* represented by $P_l m_l$ (cos θ).

$$\psi_2 = C\,P_{l_{m_l}}(\cos\theta) \text{ where } l = |m_l|, |m_l|+1, |m_l|+2, \ldots \qquad \text{...(5)}$$

The constant C can be calculated by normalising the function ψ_2.

36. *Solve the radial part of the Schrodinger equation for the hydrogen atom. Obtain the energy eigen values and radial wave function.* **(Garhwal 1991, Poorvanchal 2002, P.U. 2008)**

Ans. The *radial equation* is

$$\frac{1}{\psi_1}\frac{d}{dr}\left(r^2\frac{d\psi_1}{dr}\right)+\frac{2\,mr^2}{\hbar^2}\left(E+\frac{e^2}{4\pi\epsilon_0 r}\right) = l(l+1) \qquad \text{...(1)}$$

$$\frac{d^2\psi_1}{dr^2}+\frac{2}{r}\frac{d\psi_1}{dr}+\frac{2m}{\hbar^2}\left[E+\frac{e^2}{4\pi\epsilon_0 r}\right]\psi_1 = \frac{l(l+1)}{r^2}\psi_1$$

$$\text{or} \quad \frac{d^2\psi_1}{dr^2}+\frac{2}{r}\frac{d\psi_1}{dr}+\left[\frac{2mE}{\hbar^2}+\frac{2me^2}{4\pi\varepsilon_0\hbar^2 r}-\frac{l(l+1)}{r^2}\right]\psi_1 = 0 \qquad \text{...(2)}$$

Putting $\frac{-2mE}{\hbar^2} = \beta^2$ in Eq. (2), we get

[Note. Negative sign with energy E indicates that the electron is bound to the nucleus]

$$\frac{d^2\psi}{dr^2}+\frac{2}{r}\frac{d\psi_1}{dr}+\left[-\beta^2+\frac{2me^2}{4\pi\varepsilon_0\hbar^2 r}-\frac{l(l+1)}{r^2}\right]\psi_1 = 0$$

$$\text{or} \quad \frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d\psi_1}{dr}\right)+\left[-\beta^2+\frac{2me^2}{4\pi\varepsilon_0\hbar^2 r}-\frac{l(l+1)}{r^2}\right]\psi_1 = 0 \qquad \text{...(3)}$$

Now put $\rho = 2\,\beta\,r$ (where ρ is a dimension-less variable.) ...(4)

$$\therefore \qquad \frac{d\rho}{dr} = 2\beta$$

$$\text{Also} \qquad \frac{d\psi_1}{dr} = \frac{d\psi_1}{d\rho}\cdot\frac{d\rho}{dr} = 2\beta\frac{d\psi_1}{d\rho}$$

$$\text{or} \qquad r^2\frac{d\psi_1}{dr} = 2\beta r^2\frac{d\psi_1}{d\rho} = 2\beta\frac{\rho^2}{4\beta^2}\frac{d\psi_1}{d\rho} = \frac{\rho^2}{2\beta}\frac{d\psi_1}{d\rho}$$

Differentiating both sides of this equation w.r.t. r, we get

$$\frac{d}{dr}\left(r^2\frac{d\psi_1}{dr}\right)=\frac{d}{dr}\left(\frac{\rho^2}{2\beta}\frac{d\psi_1}{d\rho}\right) = \frac{d}{d\rho}\left(\frac{\rho^2}{2\beta}\frac{d\psi_1}{d\rho}\right)\frac{d\rho}{dr} = \frac{d}{d\rho}\left(\rho^2\frac{d\psi_1}{d\rho}\right) \qquad ...(5)$$

Substituting the value $r = \rho/2\beta$ from Eq. (4) and $\frac{d}{dr}\left(r^2\frac{d\psi_1}{dr}\right)$ from Eq. (5) in Eq. (3), we get

$$\frac{4\beta^2}{\rho^2}\frac{d}{d\rho}\left(\rho^2\frac{d\psi_1}{d\rho}\right)+\left[-\beta^2+\frac{me^2\beta}{\pi\varepsilon_0\hbar^2\rho}-\frac{4l(l+1)\beta^2}{\rho^2}\right]\psi_1 = 0$$

or
$$\frac{d^2\psi_1}{d\rho^2}+\frac{2}{\rho}\frac{d\psi_1}{d\rho}+\left[-\frac{1}{4}+\frac{me^2}{4\pi\varepsilon_0\hbar^2\rho\beta}-\frac{l(l+1)}{\rho^2}\right]\psi_1 = 0$$

or
$$\frac{d^2\psi_1}{d\rho^2}+\frac{2}{\rho}\frac{d\psi_1}{d\rho}+\left[\frac{n}{\rho}-\frac{l(l+1)}{\rho^2}-\frac{1}{4}\right]\psi_1 = 0 \qquad ...(6)$$

where
$$n = \frac{me^2}{4\pi\varepsilon_0\hbar^2\beta} \qquad ...(6)(a)$$

Asymptotic solution. For large ρ, the only terms that remain in Eq. (6) are

$$\frac{d^2\psi_1}{d\rho^2}-\frac{\psi_1}{4} = 0$$

The solution, which behaves properly at infinity, is $\psi_1 = Ae^{-\rho/2}$

With the asymptotic solution, the possible solution of Eq. (6) is

$$\psi_1 = Ae^{-\rho/2}F(\rho) \qquad ...(7)$$

where $F(\rho)$ is a polynomial of finite order in ρ.

Substitution of Eq. (7) into Eq. (6) gives

$$\frac{d^2F}{d\rho^2}+\left(\frac{2}{\rho}-1\right)\frac{dF}{d\rho}+\left[\frac{n-1}{\rho}-\frac{l(l+1)}{\rho^2}\right]F = 0 \qquad ...(8)$$

In order that the solution of Eq. (8) *remains finite* for all values of ρ, the condition is

$$n-1 = l, l+1, l+2, l+3 \text{ ...etc.}$$

or
$$n = l+1, l+2, l+3 \text{ ... etc.}$$

n is called the *principal a quantum number.*

$$l = 0, 1, 2, \ldots (n-1),$$

$$i.e., n = 1, 2, 3, 4, \ldots$$

The energy eigenvalues. From Eq. (6) (*a*), $\beta = \frac{me^2}{4\pi\varepsilon_0\hbar^2}\cdot\frac{1}{n}$

or
$$\beta^2 = \frac{m^2e^4}{16\pi^2{\varepsilon_0}^2\hbar^4}\cdot\frac{1}{n^2}$$

But
$$\beta^2 = \frac{-2mE}{\hbar^2}$$

∴
$$\frac{-2mE}{\hbar^2} = \frac{m^2e^4}{16\pi^2\varepsilon_0^2\hbar^4}\cdot\frac{1}{n^2}$$

$$E_n = -\frac{me^4}{32\pi^2\varepsilon_0^2\hbar^2}\cdot\frac{1}{n^2} \qquad ...(9)$$

Eq. (9) is in agreement with the old quantum theory and with experiment.

Radial wave function. The radial wave function of the hydrogen atom is given by Eq. (7).

$$R = Ae^{-\rho/2} F(\rho)$$

The solution of Eq. (8) will depend upon two quantum numbers n and l and $\rho = 2\beta r$.

$$\therefore \qquad R_{nl}(\rho) = A_{nl} e^{-\rho/2} F_{nl}(\rho) = A_{nl} e^{-\beta r} F_{nl}(2\beta r) \qquad \text{...(10)}$$

Now
$$\beta = \frac{me^2}{4\pi\varepsilon_0 \hbar^2} \cdot \frac{1}{n} = \frac{1}{a_0 n}$$

Here, $a_0 = \dfrac{4\pi\varepsilon_0 \hbar^2}{me^2}$ is the radius of the first Bohr orbit.

$$\therefore \qquad \mathrm{R}_{nl} = A_{nl}\, e^{-r/a_0 n} F_{nl}\left(\frac{2r}{a_0 n}\right) \qquad \text{...(11)}$$

It can be shown that the polynomial $F_{nl}\left(\dfrac{2r}{a_0 n}\right)$ is the product of $\left(\dfrac{2r}{a_0 n}\right)^t$ and the *associated Laguerre polynominal* $L_{n+l}^{2l+1}\left(\dfrac{2r}{na_0}\right)$. Thus

$$F_{nl}\left(\frac{2r}{a_0 n}\right) = \left(\frac{2r}{a_0\, n}\right)^l L_{n+l}^{2l+1}\left(\frac{2r}{na_0}\right)$$

Thus the acceptable radial wave functions are

$$R_{nl} = A_{nl}\, e^{-r/na_0} \left(\frac{2r}{na_0}\right)^l L_{n+l}^{2l+1}\left(\frac{2r}{na_0}\right) \qquad \text{...(12)}$$

The first three normalised radial wave functions are given below.

$$R_{10}(r) = \frac{2}{a_0^{3/2}} e^{-r/a_0}$$

$$R_{20}(r) = \frac{1}{(2a_0)^{3/2}}\left(2 - \frac{r}{a_0}\right) e^{-r/2a_0}$$

$$R_{21}(r) = \frac{1}{(2a_0)^{3/2}} \frac{1}{\sqrt{3}} \frac{r}{a_0} e^{-r/2a_0}$$

Complete wave function for the hydrogen atom.

The complete hydrogen atom wave functions are given by

$$\begin{aligned} \psi(r, \theta, \phi) &= \psi_1(r)\, \psi_2(\theta)\, \psi_3(\phi) \\ &= A_{nl} e^{-r/na_0} \left(\frac{2r}{n\, a_0}\right)^l L_{n+l}^{2l+1}\left(\frac{2r}{n\, a_0}\right) C_{lm_i} P_{lm_i}(\cos\theta) \frac{1}{\sqrt{2\pi}} \exp(im_l\phi) \end{aligned}$$

$$\psi(r, \theta, \phi) = N_{nlm_i} e^{-r/na_0} \left(\frac{2r}{na_0}\right)^l L_{n+l}^{2l+1}\left(\frac{2r}{na_0}\right) P_{lm_i}(\cos\theta) \exp(im_l\phi)$$

Here, $N_{nlm_l} = A_{nl}\, C_{lm_l} \dfrac{1}{\sqrt{2\pi}}$ is called the normalisation constant.

Ground state wave function.

$$\psi_{100} = \frac{1}{\sqrt{2\pi}} \cdot \frac{1}{\sqrt{2}} \frac{2}{a_0^{3/2}} e^{-r/a_0} = \frac{1}{\sqrt{\pi}\, a_0{}^{3/2}} e^{-r/a_0}$$

37. *Outline Dirac's bra and ket notation. What is Hilbert space ?*

Ans. **The Hilbert Space of State Vectors; Dirac Notation.**

(*a*) State Vectors and Their Conjugates: Dirac introduced the symbol | ψ > to denote an abstract state vector as distinct from its representation, the wave function.

The state vectors constitute a complex vector space, i.e.,

If $| \psi >, | \chi >$ are state vectors, so is $c | \psi > + c' | \chi >$ for any complex numbers c and c'.

All vectors $c | \psi >, c \neq 0$, which have the same 'direction' as $| \psi >$ correspond to the *same* physical state. The set of all such vectors constitute what is called a *ray* in the space. Therefore, we say that any physical state is represented by a ray (rather than a single vector) in the vector space.

Corresponding to every vector $| \psi >$ is defined a *conjugate vector* $\overline{| \psi >}$, for which Dirac used the notation $< \psi |$.

$$< \psi | \equiv \overline{| \psi >}, \overline{< \psi |} \equiv | \psi > \qquad ...(1)$$

In the Schrödinger representation $| \psi > \rightarrow \psi(\mathbf{x}), < \psi | \rightarrow \psi^*(\mathbf{x})$.

Thus, the state vector corresponding to $\psi_a(\mathbf{r})$ is denoted by the *ket* $| a >$.

Corresponding to every vector, $| a >$ is defined a conjugate vector $| a >^*$ for which Dirac used the notation $< a |$ which is called a *bra*.

- Kets $| a >$ and $| b >$ are *orthogonal* if $< a | b > = 0$
- The ket $| a >$ is said to be *normalized* if $< a | a > = 1$
- The *orthonormality relation* is expressed as $< a_i | a_j > = \delta_{ij}$
- The condition for an operator to be Hermitian is $< a | A | b > = < b | A | a >^*$

(*b*) *Norm and Scalar Product*: The squared length or norm of the vector $| \psi >$ is denoted by $< \psi | \psi >$. By definition, $< \psi | \psi >$ is a real non-negative number :

$$< \psi | \psi > \geq 0, < \psi | \psi > = 0 \text{ if and only if } | \psi > = 0. \qquad ...(2a)$$

More generally, when any vector $| \psi >$ and the conjugate $< \phi |$ of some vector $| \phi >$ are juxtaposed in such a way as to form a *closed* bracket, $< \phi | \psi >$, this symbol is understood to mean a single (real or complex) number which is defined to be the *scalar product* of $| \phi >$ and $| \psi >$, taken in that order.

Vectors of the type $< \psi |$ which go to make up the *first half* of such *closed brackets* are called *bra vectors* and those of the type $| \psi >$ constituting the second half are called *ket vectors*. The conjugate of a bra vector is a ket vector, and vice versa.

In the Schrödinger represention, $< \phi | \psi > \rightarrow \int \phi^*(\mathbf{x}) \psi(\mathbf{x}) \, d\tau$.

Note that $< \phi | \psi > \neq < \psi | \phi >$; instead

$$< \phi | \psi > = < \psi | \phi >^* \qquad ...(2b)$$

The scalar product of two vectors is linear in the second vector :

$$< \phi | (c | \psi > + c' | \chi >) = c < \phi | \psi > + c' < \phi | \chi > \qquad ...(2c)$$

Any vector space (in general infinite dimensional) within which a scalar product with the properties (2) is defined between any two vectors is called a *Hilbert Space*.

The space of ket vectors (or bra vectors) representing quantum mechanical states is thus a Hilbert space.

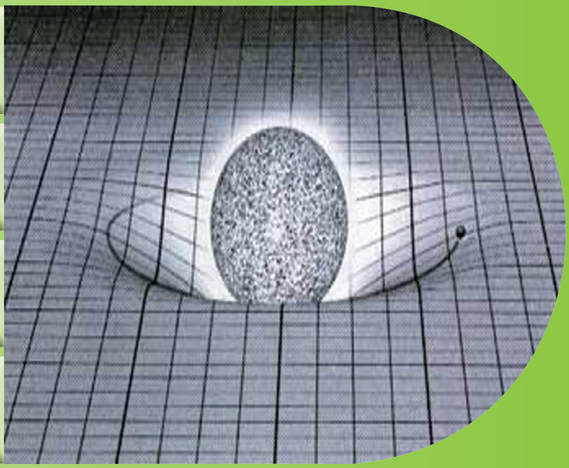
(*c*) *Basis in Hilbert Space*: Let $| \phi_1 >, | \phi_2 >$, ... be a linearly independent set of states in the Hilbert space, i.e., such that no linear combination of these vectors vanishes except when all the coefficients vanish. Let the set be also a *complete* one, in the sense that any arbitrary vector $| \psi >$ of the space can be expressed in the form

$$| \psi > = c_1 | \phi_1 > + c_2 | \phi_2 > + ... = \sum_i c_i | \phi_i > \qquad ...(3)$$

Then the set $| \phi_1 >, | \phi_2 >$, ... can be chosen as a *basis* or 'coordinate system' for the Hilbert space. The set of numbers c_1, c_2, ... in Eq. (3) are the 'components' of $| \psi >$ with respect to this basis, and it provides a *representation of* $| \psi >$.

CHAPTER

12



QUANTUM MECHANICS II

AT A GLANCE			
12.1	Postulates of Quantum Mechanics	**12.2**	Probability Current Density
12.3	The Free Particle	**12.4**	Rectangular Potential Well
12.5	The Square Well in Three Dimensions	**12.6**	Ehrenfest's Theorem
12.7	Mathematical Proof of Uncertainty Principle for One Dimensional Wave-packet	**12.8**	Wave Velocity (or Phase Velocity)

In chapter 11, we have developed Schrodinger wave equation and have found its solution for some cases of physical interest, *e.g.*, particle in a box (infinite square well potential), potential step, potential barrier, linear harmonic oscillator, the hydrogen atom and the rigid rotator. In this chapter, we shall study the following systems: the free particle, rectangular potential well, and the square well in three dimensions. We shall also give a mathematical proof of "uncertainty principle".

12.1 Postulates of Quantum Mechanics

Postulate 1. *Any state of a system with n degrees of freedom is described as completely as possible by a wave-function* $\psi\ (q_1, q_2, ...,q_n\ t)$.

The wave-function depends on the coordinates of all the particles in the system and the time. For a single particle, the *wave-function* is a function only of position $\mathbf{r}$ and time t, and is written $\psi\ (\mathbf{r}, t)$. The wave function $\psi\ (\mathbf{r}, t)$ gives the complete knowledge of the behaviour of the particle. Similarly, $\psi\ (\mathbf{r})$ gives the stationary state which in independent of time. Quantum mechanics assigns to every system a wave function ψ that contains implicitly everything that can be known about that system. The function ψ is continuous. ψ can be subjected to differentiation.

The probability of finding the particle in the space within r and $r + dr$ is $\psi^*\ \psi\ d\ r$.

Postulate 2. *Every physical observable is associated with a linear Hermitian operator.*

All the observables, that is, measurable quantities can be represented as mathematical operators.

The quantum mechanical operator for the total energy is

$$\hat{H} = \begin{cases} -\dfrac{\hbar^2}{2m}\nabla^2 + V \text{ (time independent form)} \\ i\hbar\dfrac{\partial}{\partial t} \qquad\qquad \text{(time dependent form)} \end{cases}$$

$\hat{H}$ is called the *Hamiltonian* (energy) *operator.*

(*i*) Linear Operator. An operator transforms one function into another. A well-known example is the differential operator $D \equiv d/dx$, e.g. $Dx^3 = 3x^2$

In quantum mechanics we are concerned with linear operators.

An operator α is said to be linear if

$$\alpha\,(\psi_1 + \psi_2) = \alpha\,\psi_1 + \alpha\,\psi_2 \text{ and } \alpha\,(a\,\psi) = a\,\alpha\,\psi.$$

ψ_1 and ψ_2 are arbitrary functions. a is a constant which may or may not be complex.

The operator D is linear.

Commutator. Suppose α and β are two quantum operators. The commutator of these two operators is defined to be

$$[\alpha, \beta] \equiv (\alpha\,\beta - \beta\,\alpha)$$

Suppose $\alpha = x$ and $\beta = d/dx$. Then

$$\left[\quad \frac{}{dx}\right] \equiv \left[x\frac{d}{dx} - \frac{d}{dx}x\right] \equiv \left[x\frac{d}{dx} - 1 - x\frac{d}{dx}\right] \equiv -1$$

Thus $\alpha\,\beta \neq \beta\,\alpha$

Two operators α and β are said to commute with each other when

$$\alpha\,\beta - \beta\,\alpha = 0$$

For example, if $\alpha = \partial/\partial x$ and $\beta = a\,\times$, where a is a constant, then we get $\alpha\,\beta - \beta\,\alpha = 0$. α and β commute with each other. When two operators commute, it is possible to measure both the variables associated with them simultaneously and accurately. Further, there will be no uncertainty in either of them.

If two operators do not commute, it is not possible to measure the variables associated with them simultaneously and accurately. If one of them is measured accurately, it is not possible to measure the other simultaneously with accuracy.

(*ii*) Hermitian operator. An operator α operating on any two functions ψ_1 and ψ_2 is *Hermitian*, if

$$\int \psi_1^*\,(\alpha\,\psi_2)\,d\tau = \int \psi_2\,(\alpha\,\psi_1)^*\,d\tau\,.$$

Clearly a linear combination of Hermitian operators α, β is itself a Hermitian operator.

EXAMPLE 1. *Prove that the eigen values of a Hermitian operator are real.*

Let the Hermitian operator be α and ϕ_n an eigen function belonging to the eigen value a_n. That is

$$\alpha\phi_n = a_n\,\phi_n \qquad \text{...(1)}$$

$$\therefore \qquad \int \phi^*{}_n\,\alpha\phi_n\,d\tau = a_n \int \phi^*{}_n\,\phi_n\,d\tau$$

Take the complex conjugate of (1).

$$(\alpha\,\phi_n)^* = a_n^*\ \phi^*{}_n$$

$$\therefore \qquad \int (\alpha\phi_n)^* \phi_n \, d\tau = a^*{}_n \int \phi^*{}_n \phi_n \, d\tau$$

As the operator is Hermitian

$$\int \phi_n^* \alpha\phi_n \, d\tau = \int (\alpha\phi_n)^* \phi_n \, d\tau$$

$$\therefore \qquad a_n = a_n^*$$

and the eigen value a_n is real.

EXAMPLE 2. *Prove that any two eigen functions of a Hermitian operator, belonging to different eigen values are orthogonal.*

SOL. Let the Hermitian operator be α. Let the eigen functions be ϕ_n, ϕ_m with corresponding eigen values a_n, a_m. That is

$$\alpha\,\phi_n = a_n\,\phi_n \qquad \text{...(1)}$$

$$\alpha\,\phi_m = a_m\,\phi_m \qquad \text{...(2)}$$

From (1),

$$\int \phi^*{}_m \alpha\phi_n \, d\tau = a_n \int \phi^*{}_m \phi_n \, d\tau$$

$$\int \phi^*{}_m \alpha\phi_n \, d\tau = \int \phi_n (\alpha\,\phi_m)^* \, d\tau \qquad [\because \alpha \text{ is Hermitian}]$$

$$= \int \phi_n \, a_m{}^* \, \phi^*{}_m d\tau \qquad \text{(From (2)]}$$

$$= a_m \int \phi^*{}_m \phi_n \, d\tau \qquad [\because a_m \text{ is real}]$$

$$\therefore \qquad (a_n - a_m)\int \phi^*{}_m \phi_n \, d\tau = 0$$

As $a_n \neq a_m$, we have $\int \phi^*{}_m \phi_n \, d\tau = 0.$

This indicates that ϕ_n and ϕ_m are *orthogonal functions.*

(*iii*) Simultaneous eigenfunctions; commutators

If A and B are linear operators, and ψ is a function satisfying both the equations

$$A\psi = \alpha\psi, \quad B\psi = \beta\psi, \qquad \text{...(1)}$$

then ψ is a *simultaneous* eigenfunction of A and B, belonging to the eigenvalues α and β, respectively.

For example, consider a nonrelativistic free particle, of energy $E = \frac{1}{2}mv^2$ and momentum $p = mv$. The wave function

$$\psi_p = \frac{1}{\sqrt{2\pi\hbar}} e^{(i/\hbar)px}, \qquad \text{...(2)}$$

satisfies simultaneously the equations

$$p_{op}\psi = p\psi \quad \text{and} \quad H_{op}\psi = E\psi.$$

Here, $H_{op} = p^2{}_{op}/2m$ is the Hamiltonian operator for a free particle.

Equations (1) imply that

$$B\,A\psi = B(\alpha\psi) = \alpha B\psi = \alpha\beta\psi,$$

and

$$A\,B\psi = A(\beta\psi) = \beta\,A\psi = \beta\alpha\psi.$$

By subtraction,

$$(AB - BA)\,\psi = 0. \qquad \text{...(3)}$$

This equation shows that ψ is also an eigenfunction of the operator $(AB - BA)$, belonging to the eigenvalue zero. The condition (3) is necessary in order that ψ be a simultaneous eigenfunction of A and B.

The operator in Eq. (3) is called the *commutator* of A and B and is written,

$$[A, B] = AB - BA. \quad \text{...(4)}$$

Two operators satisfying the equation

$$[A, B] = 0 \quad \text{...(5)}$$

are said to *commute.* This equation means that Eq. (3) is true for every ψ which is a member of the class of functions under consideration.

The eigenfunctions of commuting operators can always be constructed in such a way that they are simultaneous eigenfunctions.

(*iv*) Simultaneous Measurability of Observables

If two observables are simultaneously measurable in a particular state of a given system, then the state function is an eigenfunction of both the operators. Two observables are said to be *compatible,* if their operators have a common set of eigenfunctions. The following two theorems indicate the connection between compatible observables and commuting operators.

Theorem 1. *Operators having common set of eigenfunctions commute.*

Proof. Consider operators A and B with the common set of eigenfunctions ψ_i, $i = 1, 2, \ldots$ as

$$A\psi_i = a_i \psi_i, \quad \text{and} \quad B\psi_i = b_i\psi_i \quad \text{...(1)}$$

Then

$$AB\psi_i = A(b_i\psi_i) = b_iA\psi_i = a_ib_i\psi_i \quad \text{...(2)}$$

and

$$BA\psi_i = B(a_i\psi_i) = a_iB\psi_i = a_ib_i\psi_i \quad \text{...(3)}$$

Since $AB\psi_i = BA\psi_i$, A commutes with B. Hence the result.

Theorem 2. *Commuting operators have common set of eigenfunctions.*

Proof. Consider two commuting operators A and B. The eigenvalue equation for A is

$$A\psi_i = a_i\psi_i, \quad i = 1, 2, \quad \text{...(1)}$$

Operating both sides from left by B, we get

$$BA\psi_i = a_iB\psi_i$$

Since B commutes with A,

$$A(B\psi_i) = a_i(B\psi_i) \quad \text{...(2)}$$

That is, $B\psi_i$ is an eigenfunction of A with the same eigenvalue a_i. If A has only nondegenerate eigenvalues, $B\psi_i$ can differ from ψ_i only by a multiplicative constant, say b_i

$$B\psi_i = b_i\psi_i \quad \text{...(3)}$$

In other words, ψ_i is a simultaneous eigenfunction of both A and B.

Angular Momentum Operator. The angular momentum about the origin of a particle with linear momentum **p** at the position **r** is

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} \quad \text{...(1)}$$

The operator associated with linear momentum **p** is $-i\hbar\nabla$.

Thus the *angular momentum operator is*

$$\mathbf{L} = \mathbf{r} \times (-i\hbar\nabla) = -i\hbar\,\mathbf{r} \times \nabla \quad \text{...(2)}$$

Let L_x, L_y, L_z be cartesian components of **L**.

$$\boldsymbol{i}L_x + \boldsymbol{J}L_y + \boldsymbol{k}L_z = -i\hbar \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ x & y & z \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \end{vmatrix}$$

$$\therefore \quad \hat{L}_x = -i\hbar\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right)$$

$$\hat{L}_y = -i\hbar\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right)$$

$$\hat{L}_z = -i\hbar\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right) \qquad ...(3)$$

$\hat{L}_x, \hat{L}_y, \hat{L}_z$ do not commute with one another.

$$[\hat{L}_x, \hat{L}_y] = i\hbar\hat{L}_z; [\hat{L}_y, \hat{L}_z] = i\hbar\hat{L}_x; [\hat{L}_z, \hat{L}_x] = i\hbar\hat{L}_y .$$

The three components of the angular momentum operators commute with the Hamiltonian :

$$[\hat{H}, \hat{L}_x] = 0; [\hat{H}, \hat{L}_y] = 0; [\hat{H}, \hat{L}_z] = 0$$

In spherical polar coordinates,

$$\hat{L}_x = i\hbar\left(\sin\phi\frac{\partial}{\partial\theta} + \cot\theta\cos\phi\frac{\partial}{\partial\phi}\right)$$

$$\hat{L}_y = -i\hbar\left(\cos\phi\frac{\partial}{\partial\theta} - \cot\theta\sin\phi\frac{\partial}{\partial\phi}\right)$$

$$\hat{L}_z = -i\hbar\frac{\partial}{\partial\phi}$$

It can be proved that

$$\hat{L}^2 = \hat{L}_x{}^2 + \hat{L}_y{}^2 + \hat{L}_z{}^2$$

$$= -\hbar^2\left\{\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right\}$$

It can be shown that $\hat{L}^2$ commutes with each of the three components $\hat{L}_x, \hat{L}_y, \hat{L}_z$:

$$[\hat{L}^2, \hat{L}_x] = 0,\ [\hat{L}^2, \hat{L}_y] = 0,\ [\hat{L}^2, \hat{L}_z] = 0$$

Postulate 3. *The development in time of the wave function* ψ *of a system is given by the equation*

$$\hat{H}\psi = i\hbar\frac{\partial}{\partial t}\psi$$

where $\hat{H}$ is the Hamiltonian of the system.

This postulate is of course simply a statement of the time dependent Schrodinger equation. If the wave function is known at some initial time, the above equation determines ψ at any other time.

Postulate 4. *The only possible values which a measurement of the observable* (*whose operator is* α) *can yield are the eigen values* a_n *of the equation*

$$\alpha\,\psi_n = a_n\,\psi_n$$

This is true provided $\int \psi^*_n \psi_n\, d\tau < \infty$ and ψ_n is single-valued.

Example 1. *Which of the following are eigen functions of the operator* d^2/dx^2? *Give the eigen values where appropriate.* (*a*) sin *x* (*b*) cos *x* (*c*) e^x (*d*) e^{ix} (*e*) $\sin^2 x$.

Sol. Let *Q* be the operator and *f* be the function.

(*a*) $Qf = \frac{d^2}{dx^2}\sin x = -\sin x = -f$. The eigen value is -1.

(*b*) $Qf = \frac{}{dx}\cos x = -\cos x = -f$. The eigen value is -1.

(*c*) $Qf = \frac{d^2}{dx^2}e^x = e^x = f$. The eigen value is $+1$.

(*d*) $Qf = \frac{d^2}{dx^2}e^{ix} = -e^{ix} = -f$. The eigen value is -1.

(*e*) $Qf = \frac{d^2}{dx^2}\sin^2 x = 2 - 4\sin^2 x$. It is not an eigen function.

EXAMPLE 2. *Which of the following wave functions are acceptable in quantum mechanics?* (*a*) sin *x* (*b*) tan *x* (*c*) cosec *x* (*d*) cos *x* + sin *x*.

SOL. (*a*) and (*d*) are acceptable wave functions. (*b*) and (*c*) are not acceptable because (*b*) → ∞ at $x = \pi/2$ and (*c*) → ∞ at $x = 0$.

Postulate 5. *The average value of an observable a, corresponding to the operator* α, *for a system described by the wave function* ψ *is given by*

$$< a > = \frac{\int \psi^* \alpha \psi \, d\tau}{\int \psi^* \psi \, d\tau}$$

$< a >$ is called the *expectation value*.

12.2 Probability Current Density

Suppose ψ (***r***, *t*) is the state function representing a one-particle system.

The probability that the particle is in the finite volume τ is

$$P = \int_\tau \psi^* \psi \, d\tau \qquad ...(1)$$

The rate of change of this probability is

$$\frac{dP}{dt} = \int_\tau \left(\psi \frac{\partial \psi^*}{\partial t} + \psi^* \frac{\partial \psi}{\partial t} \right) d\tau \qquad ...(2)$$

Time-dependent Schrodinger equation is

$$-\frac{\hbar^2}{2m}\nabla^2 \psi + V\psi = i\hbar \frac{\partial \psi}{\partial t} \qquad ...(3)$$

Assuming *V* to be real, complex conjugate of Eq. (3) is

$$-\frac{\hbar^2}{2m}\nabla^2 \psi^* + V\psi^* = -i\hbar \frac{\partial \psi^*}{\partial t} \qquad ...(4)$$

$$\frac{\partial \psi}{\partial t} = \frac{1}{i\hbar}\left[-\frac{\hbar^2}{2m}\nabla^2 \psi + V\psi \right]; \frac{\partial \psi^*}{\partial t} = -\frac{1}{i\hbar}\left[-\frac{\hbar^2}{2m}\nabla^2 \psi^* + V\psi^* \right]$$

$$\therefore \qquad \psi \frac{\partial \psi^*}{\partial t} + \psi^* \frac{\partial \psi}{\partial t} = -\frac{i\hbar}{2m}\left[\psi \nabla^2 \psi^* - \psi^* \nabla^2 \psi \right]$$

$$= -\frac{i\hbar}{2m}\nabla.\left(\psi\nabla\psi^* - \psi^*\nabla\psi\right)$$

We define probability current density as

$$\mathbf{j} = \frac{i\hbar}{2m}\left[\psi\nabla\psi^* - \psi^*\nabla\psi\right] \qquad ...(5)$$

$$\therefore \quad \int_\tau \left(\psi\frac{\partial\psi^*}{\partial t} + \psi^*\frac{\partial\psi}{\partial t}\right)d\tau = -\int_\tau \nabla.\mathbf{j}\,d\tau$$

Eq. (2) becomes

$$\frac{dP}{dt} = -\int_\tau \nabla.\mathbf{j}\,d\tau \qquad ...(6)$$

The volume integral can be transformed to a surface integral by Gauss's theorem, *i.e.*,

$$\frac{dP}{dt} = -\oint_S \mathbf{j}.d\mathbf{S} \qquad ...(7)$$

Here S is the surface enclosing the volume τ. The direction of $d\boldsymbol{S}$ is along the outward normal. Clearly from Eq. (7), the integral of $\boldsymbol{j}$ over the surface S is the probability that the particle will cross the surface going outwards in unit time.

From Eqs. (2) and (6), $\frac{\partial}{\partial t}(\psi^*\psi) + \nabla.\boldsymbol{j} = 0$ ...(8)

This equation expresses the *conservation of probability density*. It is analogous to the equations of continuity of hydrodynamics and electrodynamics.

12.3 The Free Particle

A free particle is defined as one which is subject to no forces of any kind, and so, moves in a region of constant potential. The particle motion is confined to the x-axis only. We shall consider the potential to be zero, $V = 0$.

The time-independent Schrodinger equation becomes

$$\frac{d^2\psi}{dx^2} + \frac{2mE}{\hbar^2}\psi = 0 \qquad ...(1)$$

Define a quantity k, called the wave number, by

$$k = \sqrt{\frac{2mE}{\hbar^2}} \qquad ...(2)$$

$$\therefore \quad \frac{d^2\psi}{dx^2} + k^2\psi = 0 \qquad ...(3)$$

The solution of Eq. (3) is

$$\psi = A\,e^{ikx} + B\,e^{-ikx} \qquad ...(4)$$

Let us set $B = 0$. Then

$$\psi = A\,e^{ikx} \qquad ...(5)$$

Eq. (5) describes a particle moving in the positive x-direction.

A problem appears in normalising the wave function. We ought to have

$$\int_{-\infty}^{\infty}\psi^*\psi\,dx = A^2\int_{-\infty}^{\infty}dx = 1 \qquad ...(6)$$

$\int_{-\infty}^{\infty} dx$ is infinite. So A must be zero. This difficulty arises since we are considering an ideal case of infinite length. In practice, the particle is confined to a finite length so that normalisation is possible.

The probability density for the particle is $\psi^* \psi = A^2$, a constant independent of x. Consequently, the particle is equally likely to be found anywhere. So we have an infinite amount of uncertainty in its position, *i.e.*, $\Delta x = \infty$. According to the expression $\Delta p \, \Delta x \geq \hbar/2$, Δp = must be zero.

Let us prove this.

$$< p > = \int_{-\infty}^{\infty} \psi^* \left(-i\hbar \frac{\partial}{\partial x}\right) \psi \, dx = \hbar k \int_{-\infty}^{\infty} \psi * \psi \, dx = \hbar k$$

Notice that $< p >$ is positive, meaning that our particle has a momentum of exactly $\hbar k$ and is moving toward the right.

$$< p^2 > = \int_{-\infty}^{\infty} \psi^* \left(-\hbar^2\right) \frac{\partial^2}{\partial x^2} \psi \, dx = -\hbar^2 \left(-k^2\right) \int_{-\infty}^{\infty} \psi^* \psi \, dx$$

$$= \hbar^2 k^2$$

$$\therefore \qquad (\Delta p)^2 = < p^2 > - <p>^2 = 0 \, .$$

Thus the momentum of the particle is precisely defined.

Note that for the particle in a square well the energy was quantised. But for the free particle, there is no such restriction on the energy. Generally, bound systems (such as electrons trapped in an atom) will give rise to discrete energy levels. Unbound systems (such as two atoms that collide) will give rise to an energy continuum.

12.4 Rectangular Potential Well

A one dimensional rectangular potential well is defined by

$$V(x) = \begin{cases} 0 & (x < 0), \\ -V_0 & (0 < x < a), \\ 0 & (x > a), \end{cases} \qquad \ldots(1)$$

...(1)

V_0, the *depth of the well*, is a positive number (Fig. 12.1). When total energy of the particle is negative ($- V_0 < E < 0$), the particle is confined to the well and forms bound states. m is the mass of the particle.

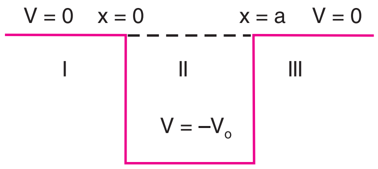


Fig. 12.1

The Schrodinger equation for the three regions are

$$\frac{d^2 \psi_1}{dx^2} + \frac{2mE}{\hbar^2} \psi_1 = 0 \text{ or } \frac{d^2 \psi_1}{dx^2} - \lambda^2 \psi_1 = 0 \quad (x < 0)$$

$$\frac{d^2 \psi_2}{dx^2} + \frac{2m(E + V_0)}{\hbar^2} \psi_2 = 0 \text{ or } \frac{d^2 \psi_2}{dx^2} + k^2 \psi_2 = 0 \quad (0 < x < a)$$

$$\frac{d^2 \psi_3}{dx^2} + \frac{2mE}{\hbar^2} \psi_3 = 0 \text{ or } \frac{d^2 \psi_3}{dx^2} - \lambda^2 \psi_3 = 0 \quad (x > a)$$

Here, $$\lambda = \sqrt{\frac{-2mE}{\hbar^2}} \qquad \text{and} \qquad k = \sqrt{\frac{2m(E+V_0)}{\hbar^2}}$$

The solutions of the Schrodinger equation outside the potential well are now $e^{\lambda x}$ and $e^{-\lambda x}$. Here λ is real. When $x \to \infty$, $e^{\lambda x}$ is not bounded. Hence it must not appear in ψ for $x > a$. Similarly, ψ must not contain $e^{-\lambda x}$ for $x < 0$.

The (unnormalised) wave functions are

$$\begin{aligned} \psi_1 &= e^{\lambda x} && (x < 0), \\ \psi_2 &= A e^{ikx} + B e^{-ikx} && (0 < x < a) \\ \psi_3 &= C e^{-\lambda x} && (x > a) \qquad \text{...(2)} \end{aligned}$$

The continuity of ψ and $d\psi/dx$ *at* $x = 0$ and $x = a$ require that

$$\psi_1 = \psi_2 \text{ at } x = 0$$

$$\frac{d\psi_1}{dx} = \frac{d\psi_2}{dx} \text{ at } x = 0$$

$$\psi_2 = \psi_3 \text{ at } x = a$$

$$\frac{d\psi_2}{dx} = \frac{d\psi_3}{dx} \text{ at } x = a \cdot$$

These conditions give four relations between the three constants A, B, C.

$$\begin{aligned} 1 &= A + B \\ \lambda &= ik(A - B) \\ C e^{-\lambda a} &= A e^{ika} + B e^{-ika} \\ -\lambda\, C e^{-\lambda a} &= ik(A e^{ika} - B e^{-ika}) \qquad \text{...(3)} \end{aligned}$$

These equations will give a unique solution, only if

$$2 \cot ka = \frac{k}{\lambda} - \frac{\lambda}{k} \qquad \text{...(4)}$$

This is a transcendental equation in k.

If this condition is satisfied, the solution is

$$\text{A} = B^* = \frac{1}{2}\left(1 - i\frac{\lambda}{k}\right), C = \frac{1}{2}\left(\frac{k}{\lambda} + \frac{\lambda}{k}\right)e^{\lambda a} \sin ka \qquad \text{...(5)}$$

To solve Eq. (4) graphically, introduce the quantities

$$\gamma = \sqrt{\frac{2mV_0 a^2}{\hbar^2}}, \alpha = \frac{ka}{\gamma} = \sqrt{1 + \frac{E}{V_0}} \qquad \text{...(6)}$$

Then Eq. (4) can be expressed as

$$\gamma\alpha = (n-1)\pi + 2\cos^{-1}\alpha, (n = 1, 2, ...) \qquad \text{...(7)}$$

Fig. 12.2 illustrates this relation graphically for the two cases $\gamma = 1, 4$. The abscissae of the points of intersection (denoted by dots) of the two graphs give the energies of the corresponding states according to the formula

$$E = -V_0(1 - \alpha^2) \qquad \text{...(8)}$$

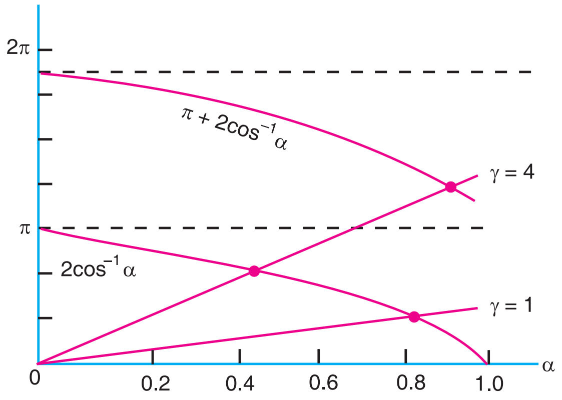


Fig. 12.2

The number of states for a given well depth is clearly the greatest integer contained in the quantity $(\gamma/\pi + 1)$. This quantity increases as V_0 is made larger. The values of E corresponding to the stationary states decreases as the well depth increases. A new level appears at zero energy each time γ assumes the value $n\,\pi$. Energy-level diagrams for the two cases of Fig. 12.2 are drawn in Fig. 12.3. n is the *quantum number* ranking the energy levels in increasing order.

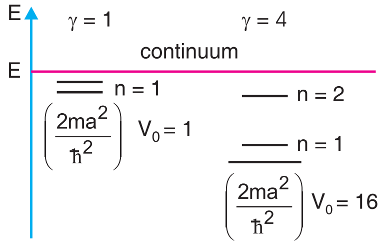

E
γ = 1
γ = 4
continuum
E
n = 1
$\left(\frac{2ma^2}{\hbar^2}\right) V_0 = 1$
n = 2
n = 1
$\left(\frac{2ma^2}{\hbar^2}\right) V_0 = 16$

Fig. 12.3

Substituting the values of A, B and C from Eq. (5) in Eq. (2),

$$\psi_1 = e^{\lambda x} \qquad (x < 0),$$

$$\psi_2 = \frac{1}{\alpha}\sin\left[k\left(x - \frac{a}{2}\right) + \frac{n\pi}{2}\right] \qquad (0 < x < a),$$

$$\psi_3 = (-1)^{n+1}\, e^{-\lambda(x-a)} \qquad (x > a) \qquad \text{...(9)}$$

ψ is zero at $x = \pm\infty$ and at $(n-1)$ points within the region $0 < x < a$. So the total number of zeros of ψ is $n + 1$. The wave function for $n = 5$ is shown in Fig. 12.4.

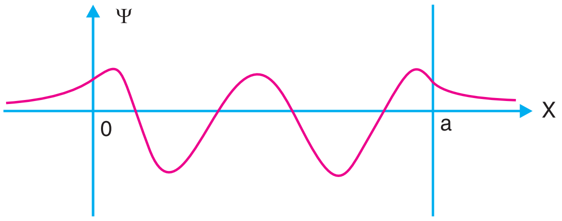


Fig. 12.4

For an *infinitely deep potential well*, the straight line of Fig. 12.2 approaches the vertical axis. The points of intersection approach the values

$$k\,a = n\,\pi \qquad (n = 1, 2, ...) \qquad \text{...(10)}$$

Hence the energy levels are

$$\frac{\hbar^2 k^2}{2m} = \frac{n^2 \pi^2 \hbar^2}{2m\,a^2} \qquad \text{...(11)}$$

For $E > 0$, well behaved solutions can be obtained for all values of E. So the energy spectrum is a continuum in this case.

12.5 The Square Well in Three Dimensions

Consider a particle of mass m which is restricted to move in a box with sides a, b and c (Fig. 12.5). The potential function $V(x, y, z)$ is having a constant value of zero in the regions given as follows :

$$V(x, y, z) = 0, \quad 0 < x < a,$$
$$V(x, y, z) = 0, \quad 0 < y < b,$$
$$V(x, y, z) = 0, \quad 0 < z < c.$$

The potential outside the box is infinite.

The Schrodinger time independent wave equation for the particle inside the box is given by

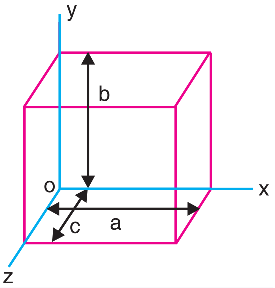


Fig. 12.5

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2} + \frac{2mE}{\hbar^2}\psi = 0 \qquad ...(1)$$

Let us solve this equation by the method of separation of variables. Assume that the function ψ can be written as a product of three functions, X, Y, and Z. Each function depends on only one of the coordinates.

$$\psi(x, y, z) = X(x)\, Y(y)\, Z(z) = X Y Z \qquad ...(2)$$

$$\frac{\partial^2 \psi}{\partial x^2} = X'' Y Z, \frac{\partial^2 \psi}{\partial y^2} = X Y'' Z, \frac{\partial^2 \psi}{\partial z^2} = X Y Z''.$$

Substituting this in Eq. (2) and dividing by XYZ, we get

$$\frac{X''}{X} + \frac{Y''}{Y} + \frac{Z''}{Z} + \frac{2mE}{\hbar^2} = 0 \qquad ...(3)$$

We notice that each term of this equation depends on a different variable and that the three variables are independent. The last term is constant. The only way for the equation to remain valid for all values of x, y, and z in our interval is for *each* term of Eq. (3) to be constant. Therefore

$$\frac{X''}{X} = -\alpha^2, \frac{Y''}{Y} = -\beta^2, \frac{Z''}{Z} = -\gamma^2$$

where α, β and γ are constants. $\alpha^2 + \beta^2 + \gamma^2 = 2m\,E/h^2$

These equations can be written as

$$\frac{d^2 X}{dx^2} + \alpha^2 X = 0 \qquad ...(4)$$

$$\frac{d^2 Y}{dy^2} + \beta^2 Y = 0 \qquad ...(5)$$

$$\frac{d^2 Z}{dz^2} + \gamma^2 Z = 0 \qquad ...(6)$$

The general solutions of (4), (5) and (6) are given by

$$X = A_1 \sin \alpha x + B_1 \cos \alpha x$$
$$Y = A_2 \sin \beta y + B_2 \cos \beta y$$
$$Z = A_3 \sin \gamma z + B_3 \cos \gamma z$$

Values of the constants $A_1, A_2, A_3, B_1, B_2, B_3$ can be found by applying boundary conditions. The boundary conditions require that the wave function vanishes at the box walls where the potential is infinite, *i.e.*,

$$\psi(0, y, z) = \psi(x, 0, z) = \psi(x, y, 0) = 0 \qquad \text{...(7)}$$

$$\psi(a, y, z) = \psi(x, b, z) = \psi(x, y, c) = 0 \qquad \text{...(8)}$$

Applying boundary conditions (7) to the above equations,

$$B_1 = B_2 = B_3 = 0$$

Applying boundary conditions (8) to the above equations

$$\sin \alpha a = 0, \textit{ i. e.}, \alpha a = n_x \pi \text{ or } \alpha = n_x\pi/a$$

$$\sin \beta b = 0, \textit{ i. e.}, \beta b = n_y \pi \text{ or } \beta = n_y\pi/b$$

$$\sin \gamma c = 0, \textit{ i. e.}, \gamma c = n_z \pi \text{ or } \gamma = n_z\pi/c$$

Here n_x, n_y, n_z are integers of which none is equal to zero.

$$\therefore \qquad X = A_1 \sin \frac{n_x \pi x}{a} \qquad \text{...(9)}$$

$$Y = A_2 \sin \frac{n_y \pi y}{b} \qquad \text{...(10)}$$

$$Z = A_3 \sin \frac{n_z \pi z}{c} \qquad \text{...(11)}$$

Substituting these values in Eq. (2), we get

$$\psi(x, y, z) = A_1 A_2 A_3 \sin \frac{n_x \pi x}{a} \sin \frac{n_y \pi y}{b} \sin \frac{n_z \pi z}{c}$$

$$= A \sin \frac{n_x \pi x}{a} \sin \frac{n_y \pi y}{b} \sin \frac{n_z \pi z}{c}$$

In the above equation A is the normalisation constant.

A can be found by using the normalisation condition

$$\int_0^a \int_0^b \int_0^c \psi \psi^* \, dx\, dy\, dz = 1 \qquad \text{...(12)}$$

$$A^2 \int_0^a \int_0^b \int_0^c \sin^2 \frac{n_x \pi x}{a} \sin^2 \frac{n_y \pi y}{b} \sin^2 \frac{n_z \pi z}{c} \, dx\, dy\, dz = 1$$

$$\therefore \qquad A^2 \frac{a}{2} \cdot \frac{b}{2} \cdot \frac{c}{2} = 1$$

or

$$A = \frac{2\sqrt{2}}{\sqrt{(abc)}} \qquad \text{...(13)}$$

The normalised wave function is

$$\psi(x, y, z) = \frac{2\sqrt{2}}{\sqrt{(abc)}} \sin \frac{n_x \pi x}{a} \sin \frac{n_y \pi y}{b} \sin \frac{n_z \pi z}{c} \qquad \text{...(14)}$$

We have $\alpha^2 + \beta^2 + \gamma^2 = 2m E/\hbar^2$

$$\frac{n_x^2 \pi^2}{a^2} + \frac{n_y^2 \pi^2}{b^2} + \frac{n_z^2 \pi^2}{c^2} = \frac{2mE}{\hbar^2}$$

$$\therefore \qquad E_{n_x n_y n_z} = \frac{\hbar^2 \pi^2}{2m} \left[\frac{n_x^2}{a^2} + \frac{n_y^2}{b^2} + \frac{n_z^2}{c^2} \right] \qquad \text{...(15)}$$

Let us examine the case in which the particle is confined in a cubical region. Then $a = b = c$. The energy levels of Eq. (15) become

$$E_{\text{cube}} = \frac{\hbar^2 \pi^2}{2ma^2}(n_x^2 + n_y^2 + n_z^2) \quad \text{...(16)}$$

The *ground state energy* value is obtained by putting

$$n_x = n_y = n_z = 1$$

$$\therefore \quad E_{111} = \frac{3\pi^2 \hbar^2}{2ma^2} \quad \text{...(17)}$$

There is only one set of quantum numbers that gives this energy state, and this level is said to be *non-degenerate*.

There are three possibilities for the first excited state :

$$\begin{aligned} n_x &= 2 \qquad & n_y = n_z = 1 \\ n_y &= 2 \qquad & n_x = n_z = 1 \\ n_z &= 2 \qquad & n_x = n_y = 1 \end{aligned}$$

$$E_{211} = E_{121} = E_{112} = \frac{3\pi^2 \hbar^2}{ma^2} \quad \text{...(18)}$$

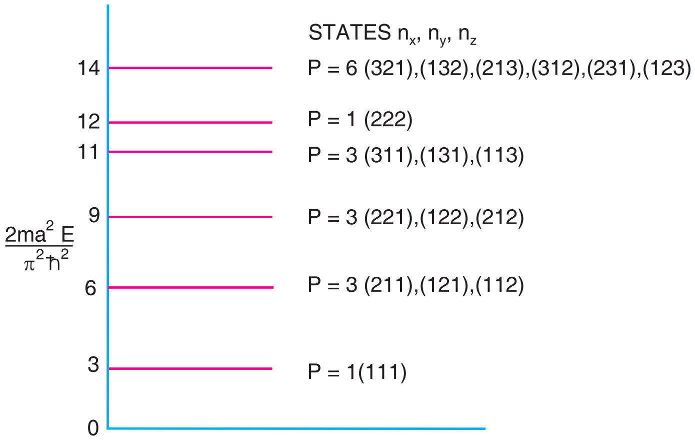


Fig. 12.6

Here there are three sets (211), (121) and (112) of the quantum numbers n_x, n_y, and n_z that will give the same energy level. That is, several distinct quantum states possess the same energy. This property is called *degeneracy*. In this particular case, the level is *triply degenerate*. Fig. 12.6 shows energy levels, degree of degeneracy and quantum numbers of a particle in a cubical box.

Eq. (15) indicates that the energy values of a particle in an infinitely deep potential well are not continuous but *discrete*. The particle has different bound states inside the well.

12.6 Ehrenfest's Theorem

According to Ehrenfest's theorem the classical equations

$$m\frac{d\mathbf{r}}{dt} = \mathbf{p} \text{ and } \frac{d\mathbf{p}}{dt} = -\nabla V$$

[$V(\boldsymbol{r})$ is the potential energy function] are also valid in quantum mechanics, provided we replace all the classical quantities by the expectation values of their corresponding quantum mechanical operators.

Paul Ehrenfest
(1880 – 1933)

Statement. *Ehrenfest theorem states that the average motion of a wave-packet agrees with the motion of the corresponding classical particle.*

In other words, it states that the Schrodinger wave equation leads to the satisfaction of classical Newton's laws of motion on the average.

Proof.

Let x be the position co-ordinate of a particle of mass m at time t.

(*a*)
$$\frac{d}{dt}<x> = \frac{d}{dt}\int \psi^* x\psi\, d^3 r$$
$$= \int \psi * x\frac{\partial\psi}{\partial t} d^3 r + \int \frac{\partial\psi *}{\partial t} x\psi\, d^3 r \qquad ...(1)$$

Time dependent Schrodinger equation is

$$\frac{\partial\psi}{\partial} = -\frac{\hbar^2}{2m}\nabla^2\psi + V\psi$$

or
$$\frac{\partial\psi}{\partial t} = -\frac{1}{i\hbar}\left\{-\frac{\hbar^2}{2m}\nabla^2\psi + V\psi\right\}$$

Complex conjugate of time-dependent Schrodinger equation is

$$-i\hbar\frac{\partial\psi^*}{\partial t} = -\frac{\hbar^2}{2m}\nabla^2\psi^* + V\psi^*$$

or
$$\frac{\partial\psi^*}{\partial t} = -\frac{1}{i\hbar}\left\{-\frac{\hbar^2}{2m}\nabla^2\psi^* + V\psi^*\right\}$$

Substituting values of $\partial\psi/\partial t$ and $\partial\psi^*/\partial t$ in Eq. (1), we get

$$\frac{d}{dt}<x> = \int \psi^* x\left\{\frac{1}{i\hbar}\left(-\frac{\hbar^2}{2m}\nabla^2\psi + V\psi\right)\right\} d^3 r$$
$$+\int\left\{-\frac{1}{i\hbar}\left(-\frac{\hbar^2}{2m}\nabla^2\psi^* + V\psi^*\right)\right\} x\psi\, d^3 r$$
$$= -\frac{\hbar}{2im}\int[\psi^* x(\nabla^2\psi) - (\nabla^2\psi^*)x\psi]\, d^3 r$$
$$\frac{d}{dt}<x> = -\frac{\hbar}{2im}\int \psi^* x(\nabla^2\psi)\, d^3 r + \frac{\hbar}{2im}\int(\nabla^2\psi^*)x\psi\, d^3 r \qquad ...(2)$$

The second integral can be integrated by parts :

$$\int(\nabla^2\psi^*)x\psi\, d^3 r = -\int(\nabla\psi^*)\cdot\nabla(x\psi)\, d^3 r + \int_A (x\psi\nabla\psi^*)_n\, dA$$

Here the integral of the normal component of $x\psi\ \nabla\psi^*$ over the infinite bounding surface A is zero because a wave packet ψ vanishes at great distances.

$$\int(\nabla^2\psi^*)x\psi\, d^3 r = -\int(\nabla\psi^*)\cdot\nabla(x\psi)\, d^3 r$$

A second partial integration, in which the surface integral again vanishes, results in

$$\int(\nabla^2\psi^*)x\psi\, d^3 r = \int\psi^*\nabla^2(x\psi)\, d^3 r \qquad ...(3)$$

Using Eq. (3), Eq. (2) can be written as

$$\frac{d}{dt}<x> = -\frac{\hbar}{2im}\int \psi^* x(\nabla^2 \psi)\, d^3 r + \frac{\hbar}{2im}\int \psi^* \nabla^2 (x\psi)\, d^3 r$$

$$= -\frac{\hbar}{2im}\int \psi^* [x\nabla^2 \psi - \nabla^2 (x\psi)]\, d^3 r$$

But $$x\nabla^2 \psi - \nabla^2 (x\psi) = -2\frac{\partial \psi}{\partial x}$$

∴ $$\frac{d}{dt}<x> = \frac{\hbar}{im}\int \psi^* \frac{\partial \psi}{\partial x} d^3 r = \frac{1}{m}\int \psi^* \left(-i\hbar\frac{\partial \psi}{\partial x}\right) d^3 r$$

or $$\frac{d}{dt}<x> = \frac{1}{m}<p_x>$$

Similarly, $$\frac{d}{dt}<y> = \frac{1}{m}<p_y> \text{ and } \frac{d}{dt}<z> = \frac{1}{m}<p_z>.$$

Combining all components,

$$\frac{d}{dt}<\mathbf{r}> = \frac{<\mathbf{p}>}{m}. \qquad ...(4)$$

(*b*) The time rate of change of *X*-component of momentum is given by,

$$\frac{d}{dt}<p_x> = \frac{d}{dt}\int \psi^* \left(-i\hbar\frac{\partial \psi}{\partial x} d^3 r\right)$$

$$= -i\left(\int \psi^* \frac{\partial}{\partial x}\frac{\partial \psi}{\partial t} d^3 r + \int \frac{\partial \psi}{\partial t}\frac{\partial \psi}{\partial x} d^3 r\right)$$

$$= -\int \psi^* \frac{\partial}{\partial x}\left(-\frac{\hbar^2}{2m}\nabla^2 \psi + V\psi\right) d^3 r + \int \left(-\frac{\hbar^2}{2m}\nabla^2 \psi^* + V\psi^*\right)\frac{\partial \psi}{\partial x} d^3 r$$

$$= -\int \psi^* \left[\frac{\partial}{\partial x}(V\psi) - V\frac{\partial \psi}{\partial x}\right] d^3 r = -\int \psi^* \frac{\partial V}{\partial x}\psi\, d^3 r = \left\langle -\frac{\partial V}{\partial x}\right\rangle$$

Similarly, $\frac{d}{dt}<p_y> = \left\langle -\frac{\partial V}{\partial y}\right\rangle$ and $\frac{d}{dt}<p_z> = \left\langle -\frac{\partial V}{\partial z}\right\rangle$

Combining all components,

$$\frac{d}{dt}<\mathbf{p}> = <-\nabla V> \qquad ...(5)$$

Equations (4) and (5) are analogous to the classical equations of motion :

$$\frac{d\mathbf{r}}{dt} = \frac{\mathbf{p}}{m} \text{ and } \frac{d\mathbf{p}}{dt} = -\nabla V$$

This proves Ehrenfest's theorem.

12.7 Mathematical Proof of Uncertainty Principle for One Dimensional Wave-packet

The uncertainty in position of a particle along *X*-axis is defined as

$$\Delta x = <(x - <x>)^2>^{1/2} \qquad ...(1)$$

Here, $<x>$ is the expectation value of x.

$\psi^*(x)\,\psi(x)\,dx$ is the probability that the position of the particle lies between x and $x + dx$. So the average or expectation of x, for normalised wave-function $\psi(x)$, can be written as

$$<x> \;=\; \int \psi^*(x)\, x\, \psi(x)\, dx \qquad ...(2)$$

Similarly, the uncertainty in momentum is defined as

$$\Delta p \;=\; <(p-<p>)^2>^{1/2} \qquad ...(3)$$

Here, the expectation value of momentum is defined as

$$<p> \;=\; \int \psi^*(x)\left(-i\hbar\frac{\partial}{\partial x}\right)\psi(x)\, dx \qquad ...(4)$$

Let us choose the normalised wave function $\psi(x)$ such that it is initially centred at $x = 0$ and has zero average momentum, *i.e.*,

$$<x> \;=\; <p> = 0 \qquad ...(5)$$

Now, consider the integral

$$i\hbar\int \psi^* \frac{d}{dx}(x\psi)\, dx$$

Integrating by parts, we have

$$\begin{aligned} i\hbar\int_{-\infty}^{\infty} \psi^* \frac{d}{dx}(x\psi)\, dx &= i\hbar[\psi x \psi^*]_{-\infty}^{\infty} - i\hbar\int_{-\infty}^{\infty} x\psi\frac{d\psi^*}{dx}\, dx \\ &= 0 - i\hbar\int_{-\infty}^{\infty} x\psi\frac{d\psi^*}{dx}\, dx \end{aligned}$$

($\because\ \psi\psi^* = 0$, for $x = \pm\infty$)

or

$$\begin{aligned} i\hbar\int \frac{d\psi^*}{dx}x\psi\, dx &= -i\hbar\int \psi^*\frac{d}{dx}(x\psi)\, dx \\ &= -i\hbar\left[\int \psi^*\frac{d\psi}{dx}x\, dx + \int \psi^*\psi\, dx\right] \end{aligned}$$

But $\int \psi^*\psi\, dx = 1$ for normalised wave function ψ.

$$\therefore \qquad i\hbar\int \frac{d\psi^*}{dx}x\psi\, dx \;=\; -i\hbar\int \psi^*\frac{d\psi}{dx}x\, dx - i\hbar$$

or

$$i\hbar\int \frac{d\psi^*}{dx}x\psi\, dx + i\hbar\int \psi^*\frac{d\psi}{dx}x\, dx = -i\hbar \qquad ...(6)$$

Equating imaginary parts of both sides and then squaring the modulus of both the sides, we get

$$4\left|\text{imaginary part of}\int i\hbar\frac{d\psi^*}{dx}x\psi\, dx\right|^2 \;=\; \hbar^2 \qquad ...(7)$$

But $\left|\int i\hbar\frac{d\psi^*}{dx}x\psi\, dx\right| \geq \left|\text{Imaginary part of}\int i\hbar\frac{d\psi^*}{dx}x\psi\, dx\right.$

since the magnitude of imaginary part of any complex number can never be greater than the magnitude of the number itself.

Hence Eq. (7) can be written as

$$4\left|\int i\hbar\frac{d\psi^*}{dx}x\psi\, dx\right|^2 \;\geq\; \hbar^2 \qquad ...(8)$$

By the inequality of Schwarz we have

$$4\left|\int i\hbar\frac{d\psi^*}{dx}x\psi\,dx\right|^2 \le 4\int i\hbar\frac{d\psi^*}{dx}\left(-i\hbar\frac{d\psi}{dx}\right)dx\int x\psi x\psi^*\,dx$$

Therefore Eq. (8) can be written as

$$4\int i\hbar\frac{d\psi^*}{dx}\left(-i\hbar\frac{d\psi}{dx}\right)dx\int \psi^* x^2\psi\,dx \ge \hbar^2$$

or $$\int\left|-i\hbar\frac{d\psi}{dx}\right|^2 dx\int \psi^* x^2\psi\,dx \ge \frac{\hbar^2}{4}$$

or $$<p^2>.<x^2> \ge \frac{\hbar^2}{4} \qquad ...(9)$$

$$\left(\because <p^2> = \int\left|-i\hbar\frac{d\psi}{dx}\right|^2 dx \text{ and } <x^2> = \int \psi^* x^2\psi\,dx\right)$$

Applying conditions expressed by Eq. (5) to Eqs. (1) and (3), we get

$$(\Delta x)^2 = <x^2> \text{ and } (\Delta p)^2 = <p^2> \qquad ...(10)$$

Using Eq. (10), Eq (9) gives $(\Delta x)^2(\Delta p)^2 \ge \frac{\hbar^2}{4}$

$$\therefore \qquad \Delta x\,\Delta p \ge \frac{\hbar}{2} \qquad ...(11)$$

12.8 Wave Velocity (or Phase Velocity)

The phase velocity u of a monochromatic wave is the velocity with which a definite phase of the wave, such as its crest or trough, is propagated in a medium.

The equation of motion of a plane wave of frequency ν and wavelength λ moving in the x-direction is

$$y = a\sin(\omega t - kx)$$

Here, a is the amplitude, ω (= 2 π ν) is the angular frequency and k (= 2 π/λ) is the propagation constant of the wave. This wave moves with a phase velocity

$$u = \nu\lambda = \omega/k$$

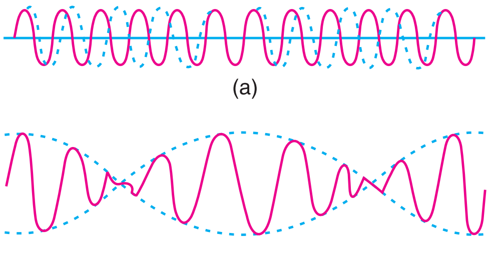


Fig. 12.7

Group Velocity. It is possible to form a concentrated *wave packet* by taking a sum of a large number of plane waves with slightly different wavelengths and frequencies. When such a group travels in a dispersive medium, the phase velocities of its different components are different. The observed velocity is, however, the velocity with which the maximum amplitude of the group advances in the medium. This is called the "group velocity". Thus, *the group velocity is the velocity with which the energy in the group is transmitted*. The individual waves travel "inside" the group with their phase velocities.

Relation between the Group Velocity (w) and the Phase Velocity (u).

Consider the superposition of two waves of equal amplitude a (Fig. 12.7). The two waves have slightly different angular frequencies ω_1 and ω_2 and propagation constants k_1 and k_2.

Their separate displacements are given by

$$y_1 = a \sin(\omega_1 t - k_1 x) \qquad \text{...(1)}$$

and

$$y_2 = a \sin(\omega_2 t - k_2 x) \qquad \text{...(2)}$$

The resultant wave obtained by the superposition of the two waves is $y = y_1 + y_2$

$$= a [\sin(\omega_1 t - k_1 x) + \sin(\omega_2 t - k_2 x)]$$

$$= 2a \cos\left[\frac{(\omega_1 - \omega_2)t}{2} - \frac{(k_1 - k_2)x}{2}\right] \sin\left[\frac{(\omega_1 + \omega_2)t}{2} - \frac{(k_1 + k_2)x}{2}\right]$$

The resultant wave has the angular frequency $(\omega_1 + \omega_2) / 2$.

Its amplitude is

$$A = 2a \cos\left[\frac{(\omega_1 - \omega_2)t}{2} - \frac{(k_1 - k_2)x}{2}\right]$$

Thus the amplitude of the wave group is modulated both in space and time by a very-slowly varying envelope of frequency $(\omega_1 - \omega_2)/2$, and propagation constant $(k_1 - k_2)/2$. This envelope is represented by the dotted curves in Fig. 12.7 (*b*). The velocity with which this envelope moves, *i.e.*, the velocity of the maximum amplitude of the group is given by

$$w = \frac{\omega_1 - \omega_2}{k_1 - k_2} = \frac{\delta\omega}{\delta k}$$

In the limit as $\delta k \to 0$, the group velocity becomes

$$w = \frac{d\omega}{dk} \qquad \text{...(3)}$$

$$= \frac{d}{dk}(ku) \qquad \because u = \frac{\omega}{k}$$

$$= u + k\frac{du}{dk}$$

$$= u + \frac{2\pi}{\lambda}\frac{du}{d(2\pi/\lambda)} \qquad \because k = \frac{2\pi}{\lambda}$$

$$= u + \frac{1}{\lambda}\frac{du}{d(1/\lambda)}$$

But

$$d\left(\frac{1}{\lambda}\right) = -\frac{1}{\lambda^2}d\lambda$$

$\therefore$

$$w = u - \lambda\frac{du}{d\lambda} \qquad \text{...(4)}$$

This is the relation between group velocity (*w*) and wave (phase) velocity (*u*) in a dispersive medium.

CHAPTER

13

QUANTUM MECHANICS III

AT A GLANCE

13.1	Normalized Wave Functions of the Hydrogen Atom	**13.2**	Expression for the Energy of the Electron of the Hydrogen Atom in the Ground State
13.3	Significance of Various Quantum Numbers	**13.4**	Electron Probability Density
13.5	Radiative Transitions	**13.6**	Selection Rules

13.1 Normalized Wave Functions of the Hydrogen Atom

In chapter 11 (Sec. 11.14), we have developed the quantum-mechanical theory of the hydrogen atom. The following table gives the normalized wave functions of the hydrogen atom for $n = 1$.

n	l	m_l	$\psi_1(r)$	$\psi_2(\theta)$	$\psi_3(\phi)$	$\psi(r, \theta, \phi)$
1	0	0	$\frac{2}{a_0^{3/2}} e^{-r/a_0}$	$\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2\pi}}$	$\frac{1}{\sqrt{\pi}\, a_0^{3/2}} e^{-r/a_0}$

The ground state wave function is given by

$$\psi_{1,0,0}(r, \theta, \phi) = \left(\frac{1}{\pi a_0^3}\right)^{1/2} \exp\left(-\frac{r}{a_0}\right) \qquad \text{...(1)}$$

where a_0 is the Bohr radius. This wave function depends only on r and not on θ and ϕ.

The four wave functions corresponding to the first excited state, as specified by the n, l, m_l values, are

$$\psi_{2,0,0}(r, \theta, \phi) = \frac{1}{4\left(2\pi a_0^3\right)^{1/2}}\left(2 - \frac{r}{a_0}\right) \exp\left(-\frac{r}{2a_0}\right) \qquad \text{...(2)}$$

$$\psi_{2,1,1}(r,\theta,\phi) = \frac{1}{8\left(\pi a_0^3\right)^{1/2}}\frac{r}{a_0}\exp\left(-\frac{r}{2a_0}\right)\sin\theta\exp(i\phi) \qquad ...(3)$$

$$\psi_{2,1,0}(r,\theta,\phi) = \frac{1}{4\left(2\pi a_0^3\right)^{1/2}}\frac{r}{a_0}\exp\left(-\frac{r}{2a_0}\right)\cos\theta \qquad ...(4)$$

$$\psi_{2,1,-1}(r,\theta,\phi) = \frac{1}{8\left(\pi a_0^3\right)^{1/2}}\frac{r}{a_0}\exp\left(-\frac{r}{2a_0}\right)\sin\theta\exp(-i\phi) \qquad ...(5)$$

We note that $\psi_{2,0,0}(r, \theta, \phi)$ is spherically symmetric, that is, depends only on r. These wave functions are normalized.

13.2 Expression for the Energy of the Electron of the Hydrogen Atom in the Ground State

The ground state normalized wave function is

$$\psi_1(r) = \frac{2}{a_0^{3/2}}e^{-r/a_0} \qquad ...(1)$$

The radial equation is

$$\frac{1}{\psi_1}\frac{d}{dr}\left(r^2\frac{d\psi_1}{dr}\right)+\frac{8\pi^2mr^2}{h^2}\left(E+\frac{e^2}{4\pi\varepsilon_0 r}\right)=l(l+1)$$

or

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d\psi_1}{dr}\right)+\left[\frac{8\pi^2m}{h^2}\left(E+\frac{e^2}{4\pi\varepsilon_0 r}\right)-\frac{l(l+1)}{r^2}\right]\psi_1 = 0 \qquad ...(2)$$

$$\frac{d\psi_1}{dr} = -\left(\frac{2}{a_0^{5/2}}\right)e^{-r/a_0}$$

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d\psi_1}{dr}\right) = \left[\frac{2}{a_0^{7/2}}-\frac{4}{ra_0^{5/2}}\right]e^{-r/a_0}$$

For the ground state, $l = 0$ and $E = E_1$.

Substituting these values in Eq. (2), we get

$$\left[\left(\frac{2}{a_0^{7/2}}+\frac{16\pi^2mE_1}{h^2a_0^{3/2}}\right)+\left(\frac{4\pi me^2}{\varepsilon_0h^2a_0^{3/2}}-\frac{4}{a_0^{5/2}}\right)\frac{1}{r}\right]e^{-r/a_0} = 0$$

Each parenthesis must equal 0 for the entire equation to equal 0.

$$\frac{4\pi me^2}{\varepsilon_0h^2a_0^{3/2}}-\frac{4}{a_0^{5/2}} = 0$$

$$\therefore \qquad a_0 = \frac{\varepsilon_0h^2}{\pi me^2} \qquad ...(3)$$

$$\frac{2}{a_0^{7/2}}+\frac{16\pi^2mE_1}{h^2a_0^{3/2}} = 0$$

or $$E_1 = -\frac{h^2}{8\pi^2 m a_0^2} = -\frac{me^4}{8\varepsilon_0^2 h^2} \qquad \text{...(4)}$$

This is the expression for the lowest energy state for the electron of the hydrogen atom.

13.3 Significance of Various Quantum Numbers

The wave functions of the hydrogen atom are characterized by three quantum numbers. Let us consider the quantum mechanical interpretation of n, l, and m_l.

1. The principal quantum number (*n*)

The radial wave equation for the hydrogen atom can be solved only when E has a continuous range of positive values, or a discrete set of *negative* values (which correspond to *bound* states), given by

$$E = -\frac{\mu e^4}{8\varepsilon_0^2 h^2}\left(\frac{1}{n^2}\right), \quad n = 1, 2, 3...$$

Thus n describes the quantisation of electron energy in the hydrogen atom.

2. The orbital quantum number (*l*)

The total energy E of the electron includes the kinetic energy because of the radial motion, kinetic energy because of the orbital motion and also the potential energy.

The total energy of the electron is

$$E = K_{radial} + K_{orbital} - \frac{e^2}{4\pi\varepsilon_0 r} \qquad \text{...(1)}$$

The radial equation is

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d\psi_1}{dr}\right) + \left[\frac{8\pi^2 m}{h^2}\left(E + \frac{e^2}{4\pi\varepsilon_0 r}\right) - \frac{l(l+1)}{r^2}\right]\psi_1 = 0 \qquad \text{...(2)}$$

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2\frac{d\psi_1}{dr}\right) + \frac{8\pi^2 m}{h^2}\left[K_{radial} + K_{orbital} - \frac{h^2 l(l+1)}{8\pi^2 m r^2}\right]\psi_1 = 0 \qquad \text{...(3)}$$

The equation is brought completely dependent on the radial motion only if we put

$$K_{orbital} - \frac{h^2 l(l+1)}{8\pi^2 m r^2} = 0$$

$$K_{orbital} = \frac{h^2 l(l+1)}{8\pi^2 m r^2} = \frac{1}{2} m v_{orbital}^2$$

Here, $v_{orbital}$ is the orbital velocity of the electron.

$$\therefore \quad l\,(l+1)\left(\frac{h}{2\pi}\right)^2 = (m v_{orbital}\, r)^2 = L^2$$

Here, $L = m v_{orbital}\, r$ = angular momentum of the electron in its orbital motion.

This means that the electron in the hydrogen atom has an angular momentum given by

$$L = \sqrt{[l(l+1)]}\,\frac{h}{2\pi}$$

where $l = 0, 1, 2, ...(n-1)$. Thus, like energy, the angular momentum is also quantised, and this quantisation is described by the orbital quantum number l. The orbital quantum number l determines the *magnitude* L of the electron's angular momentum $\boldsymbol{L}$.

3. Magnetic orbital quantum number (m_l)

An electron revolving about a nucleus (Fig. 13.1) is a current loop and has a magnetic field. Hence an atomic electron interacts with an external magnetic field ***B***. The electron orbital angular momentum vector is ***L***. The magnetic quantum number m_l specifies the direction of ***L*** by determining the component of ***L*** in the field direction. The *z*-axis is taken parallel to ***B***. θ is the angle between ***L*** and the *z*-axis. The component of ***L*** in the *z*-direction is

$$L_z = m_l\,\hbar\,.$$

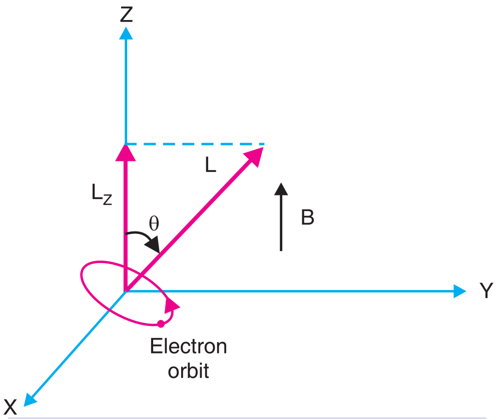


Fig. 13.1

Since, for a given l, there are $(2l + 1)$ possible values of m_l (= 0, ± 1, ± 2, ..., ± l), the angle θ can assume $(2l + 1)$ discrete values. In other words, the angular momentum vector ***L*** can have $(2l + 1)$ discrete orientations with respect to the magnetic field. This quantisation of the orientation of atoms in space is known as '*space quantisation*'.

The space quantisation of the orbital angular momentum vector ***L*** corresponding to $l = 2$ is shown in Fig. 13.2.

The uncertainty principle prohibits the angular-momentum vector ***L*** from having a definite direction in space. The simultaneous quantization of |***L***| and L_z can be described graphically by means of the vector model. According to the vector model, the angular-momentum vector ***L*** precesses around the *z*-axis, tracing out a cone. The angle between the vector ***L*** and the *z*-axis remains constant. From Fig. 13.3,

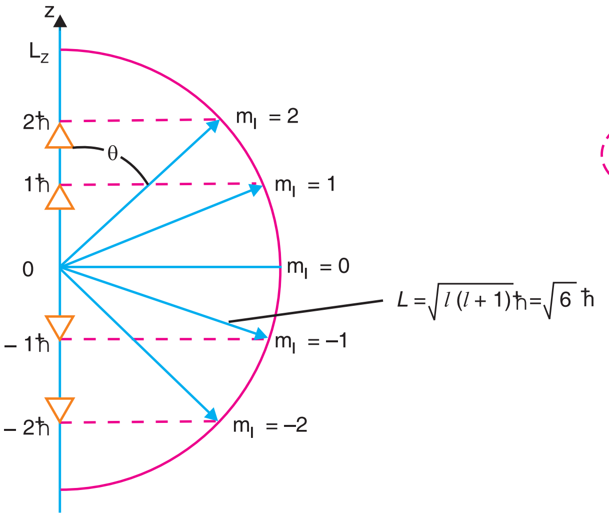


Fig. 13.2

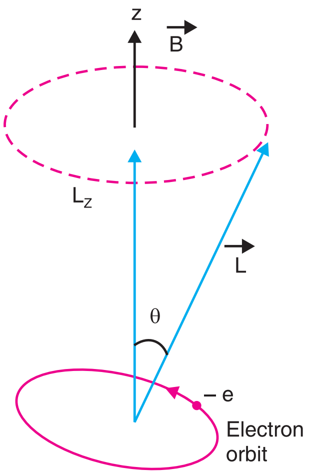


Fig. 13.3

$$\cos\theta = \frac{L_z}{|\mathbf{L}|} = \frac{m_l\hbar}{\sqrt{l(l+1)\hbar^2}} = \frac{m_l}{\sqrt{l(l+1)}}\,.$$

The smallest possible angle occurs for $m_l = l$, which gives $\cos\theta = 1\big/\sqrt{l(l+1)}$; the largest angle occurs for $m_l = 0$, which gives $\cos\theta = 0$ and $\theta = 90°$. The average values of L_x and L_y are zero, but the instantaneous values of L_x and L_y oscillate around zero. Because of the precession, the instantaneous values of L_x and L_y change quickly—successive measurements give different results, and values of L_x and L_y are not well defined. Only the average values of L_x and L_y are well defined—both averages are zero. But L_z always has the specific value $m_l\hbar$.

13.4 Electron Probability Density

The *probability density* $P(r, \theta, \phi)$ of the electron at the point (r, θ, ϕ) is proportional to $|\psi|^2$. The probability of finding the electron in a small volume dV is

$$P\,(r, \theta, \phi)\,dV = |\psi(r, \theta, \phi)|^2\,dV$$

In spherical polar coordinates

$$dV = r^2 \sin\theta\,dr\,d\theta\,d\phi$$

Thus, we have

$$P(r, \theta, \phi)\,dV = |\psi\,(r, \theta, \phi|^2\,r^2 \sin\theta\,dr\,d\theta\,d\phi \qquad ...(1)$$

The probability $P(r)\,dr$ of finding the electron between r and $(r + dr)$, irrespective of the coordinates θ, ϕ, is obtained by integrating Eq. (1) over the coordinates θ, ϕ. Thus,

$$P(r)\,dr = \int_0^{2\pi} d\phi \int_0^{\pi} d\theta \sin\theta\,|\psi(r,\theta,\phi)|^2\,r^2\,dr \qquad ...(2)$$

For the ground state of the hydrogen atom, we have

$$|\psi(r, \theta, \phi)|^2 = \frac{1}{\pi a_0^3}\exp\left(-\frac{2r}{a_0}\right) \qquad ...(3)$$

Substituting this in Eq. (2), we get

$$\begin{aligned} P(r)\,dr &= \int_0^{2\pi} d\phi \int_0^{\pi} d\theta \sin\theta \frac{1}{\pi a_0^3}\exp\left(-\frac{2r}{a_0}\right) r^2\,dr \qquad ...(4) \\ &= 2\pi \int_0^{\pi} \sin\theta\,d\theta \frac{1}{\pi a_0^3}\exp\left(-\frac{2r}{a_0}\right) r^2\,dr \\ &= [2\pi(-\cos\theta)]\Big|_0^{\pi}\ \frac{1}{\pi a_0^3}\exp\left(-\frac{2r}{a_0}\right) r^2\,dr \\ &= \left[\frac{1}{\pi a_0^3}\exp\left(-\frac{2r}{a_0}\right)\right] 4\pi r^2\,dr \qquad ...(5) \end{aligned}$$

Here, $4\pi r^2\,dr$ is the volume of a spherical shell of radius r and thickness dr. Hence, Eq. (5) is the probability density times the volume element, and can be written as

$$P(r)\,dr = |\psi_{1,0,0}|^2\,4\pi r^2\,dr. \qquad ...(6)$$

In Fig. 13.4, $P(r)\,dr$ has been plotted against r for two states of hydrogen atom corresponding to $n = 1$ and $n = 2$ ($l = 0$ in both cases). It is seen that the principal maxima of the curves closely agree with the radii of the Bohr orbits ($r = a_0$, $4a_0$, ...).

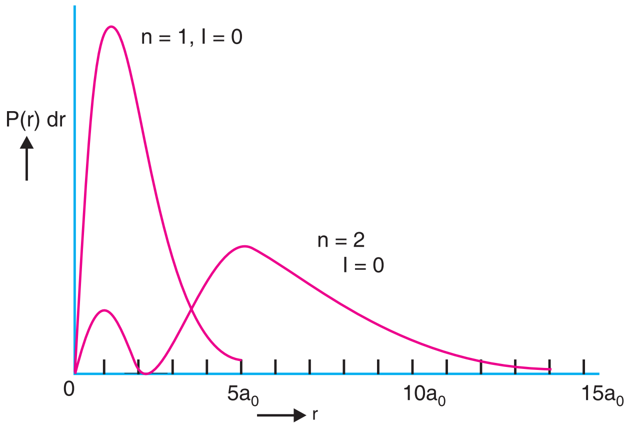


Fig. 13.4

Thus, whereas the electron is most likely to be found at the locations of Bohr orbits, it has a finite probability of being found elsewhere.

The radial probability density $P(r)$, from Eq. (5), is

$$P(r) = \frac{4}{a_0^3}\exp\left(-\frac{2r}{a_0}\right)r^2 \qquad \text{...(7)}$$

A plot of radial distribution function $P(r)$ against r is shown in Fig. 13.5

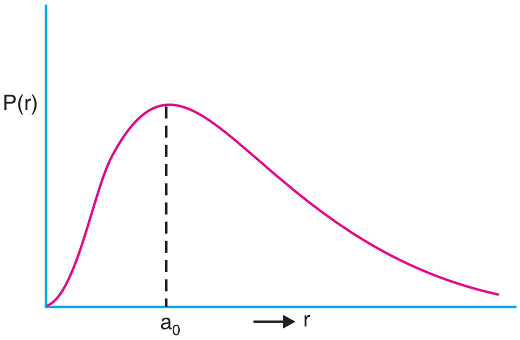


Fig. 13.5

The most probable distance of the electron from the nucleus, *i.e.*, the value of r at which $P(r)$ is maximum, is obtained by equating $\frac{dP(r)}{dr}$ to zero. That is

$$\frac{d}{dr}\left[\frac{4}{a_0^3}(e^{-2r/a_0})\,r^2\right] = 0$$

$$\text{or} \quad \frac{4}{a_0^3}\left[e^{-2r/a_0}(2r)+r^2e^{-2r/a_0}(-2/a_0)\right] = 0$$

$$\text{or} \qquad 2r + r^2(-2/a_0) = 0$$

$$\text{or} \qquad r = a_0.$$

Thus the most probable distance of the electron from the nucleus in the normal state of hydrogen atom is equal to the Bohr's radius.

EXAMPLE 1. *Calculate the mean value of r for an electron in the ground state of the hydrogen atom.*

SOL. The probability for finding the electron in a small volume dV is $|\psi_{100}|^2\,dV$. Consider a volume in the shape of a thin spherical shell of radius r and thickness dr (Fig. 13.6). For this volume, $dV = 4\pi r^2\,dr$. Hence

$$\left(\begin{array}{l}\text{probability for}\\ \text{finding electron}\\ \text{in interval from}\\ r \text{ to } r+dr\end{array}\right) = 4\pi|\psi_{100}|^2\,r^3\,dr \qquad \text{...(1)}$$

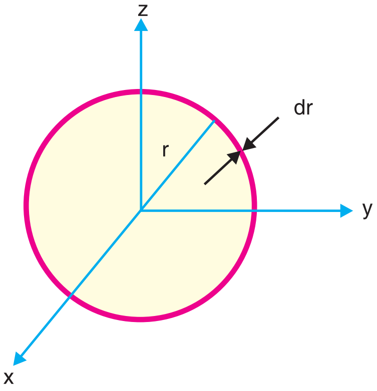


Fig. 13.6

This can also be interpreted as the probability for finding the value of r in the interval dr. To obtain the mean value of r, we must multiply this probability by r and integrate over all values of r:

$$<r> = \int_0^\infty 4\pi|\psi_{100}|^2\,r^3\,dr \qquad \text{...(2)}$$

$$= \int_0^\infty \frac{4}{a_0^3}e^{-2r/a_0}\,r^3\,dr \qquad \text{...(3)}$$

We have the standard integral,

$$\int_0^\infty e^{-\beta r}\,r^n\,dr = \frac{n!}{\beta^{n+1}} \quad \text{for } n = 0, 1, 2,... \qquad \text{...(4)}$$

$$\therefore \qquad < r > = \frac{4}{a_0^3}\frac{3!}{(2/a_0)^4} = \frac{3}{2}a_0$$

Thus the *average* value of r for a $1s$ electron is $1.5\ a_0$.

EXAMPLE 2. *Calculate the average value of* $1/r$ *for a* $1s$ *electron in the hydrogen atom.*

SOL. The expectation value of $1/r$ is

$$< \frac{1}{r} > = \int_0^\infty \left(\frac{1}{r}\right)|\psi|^2\, dV \qquad ...(1)$$

The wave function of a $1s$ electron is

$$\psi_{1,\,0,\,0} = \frac{1}{\sqrt{(\pi a_0^3)}}e^{-r/a_0}$$

$$\therefore \qquad |\psi|^2 = \frac{1}{\pi a_0^3}e^{-2r/a_0}$$

$$dV = r^2 \sin\theta\, dr\, d\theta\, d\phi.$$

Substituting these values in Eq. (1),

$$< \frac{1}{r} > = \frac{1}{(\pi a_0^3)}\int_0^\infty r(e^{-2r/a_0})\,dr \int_0^\pi \sin\theta\, d\theta \int_0^{2\pi} d\phi$$

$$= \frac{1}{\pi a_0^3}\left(\frac{a_0^2}{4}\right)(2)(2\pi)$$

$$\therefore \qquad < \frac{1}{r} > = \frac{1}{a_0}$$

13.5 Radiative Transitions

We start with the assumption that an atomic electron whose average position relative to the nucleus of the atom remains constant in time does not radiate. But if this position oscillates with time, electromagnetic waves are emitted. The frequency of these waves is the same as that of the oscillation.

The time-dependent wave function ψ of an electron is the product of a time-independent wave function ψ_0 and a time-dependent function $e^{-(iE/\hbar)t}$. The frequency of the time-varying function is $\nu = E/\hbar$.

$$\therefore \qquad \psi = \psi_0\, e^{-(iE/\hbar)t}. \qquad ...(1)$$

The expectation value $< x >$ of the position of the electron is

$$< x > = \int_{-\infty}^{\infty} x\psi^*\psi\, dx = \int_{-\infty}^{\infty} x\psi_0 {}^*\psi_0\, dx \qquad ...(2)$$

The expectation value $< x >$ is constant in time since ψ_0 and Ψ_0^* are functions of position only. No radiation results in this case when the atom is in the specific quantum state ψ, since the electron does not oscillate.

Thus, Bohr's hypothesis of non-radiating energy levels is automatically realized in the behaviour of the stationary states.

Let ψ_i and ψ_f be two states in which the atomic electron can exist. Let E_i and E_f be the respective energies of these states, with $E_i > E_f$. Let the electron be in the state ψ_f initially when it gets excited to the state ψ_i from where it then undergoes a transition to the lower state ψ_f. The wave function ψ of an electron capable of existing in both states ψ_i and ψ_f may be written as

$$\psi = a\,\psi_f + b\,\psi_i$$

Here a^*a and b^*b are the relative probabilities of the electron being in the states ψ_f and ψ_i respectively. Just before the excitation and immediately after the transition, $a = 1$, $b = 0$. When the electron is in the excited state, $a = 0$ and $b = 1$. In between, a and b have nonvanishing values. Always, $a^*a + b^*b = 1$.

Then, $< x >$ is given by

$$
\begin{aligned}
< x > &= \int_{-\infty}^{\infty} x(a^*\,\psi_f^* + b^*\,\psi_i^*)(a\psi_f + b\psi_i)\,dx \\
&= \int_{-\infty}^{\infty} \left[x(a^2\,\psi_f^*\,\psi_f + b*a\psi_i^*\psi_f + a*b\psi_f^*\psi_i + b^2\psi_i^*\psi_i)\right]dx
\end{aligned}
\qquad ...(3)
$$

Here, $a^*a = a^2$ and $b^*b = b^2$.

The integrals involving the first and the last terms in parenthesis are constant in time. The integrals involving the second and third terms are periodic in time with the frequency $(E_i - E_f)/h$. These result in electromagnetic radiation of the same frequency when the electron undergoes a transition between the states specified by ψ_i and ψ_f.

Let us expand Eq. (3) using Eq. (1).

$$
\begin{aligned}
< x > = {} & a^2 \int_{-\infty}^{\infty} x\psi_{of}^*\,\psi_{of}\,dx + b*a \int_{-\infty}^{\infty} x\psi_{oi}^*\,e^{+(iE_i/\hbar)t}\,\psi_{of}\,e^{-(iE_f/\hbar)t}\,dx \\
& + a*b \int_{-\infty}^{\infty} x\psi_{of}^*\,e^{+(iE_f/\hbar)t}\,\psi_{oi}\,e^{-(iE_i/\hbar)t}\,dx \\
& + b^2 \int_{-\infty}^{\infty} x\psi_{oi}^*\,\psi_{oi}\,dx
\end{aligned}
\qquad ...(4)
$$

Here, $\psi_{of}^*\,\psi_{oi} = \psi_{oi}^*\,\psi_{of}$ and $a*b = b*a$.

The integrals involving the second and third terms are the time-varying terms. They can be combined into the single term

$$
\begin{aligned}
& a*b \int_{-\infty}^{\infty} x\psi_{of}^*\psi_{oi}\left[e^{(i/\hbar)(E_i - E_f)t} + e^{-(i/\hbar)(E_i - E_f)t}\right]dx \\
&= 2a*b\cos\left(\frac{E_i - E_f}{\hbar}\right)t \int_{-\infty}^{\infty} x\psi_{of}^*\,\psi_{oi}\,dx
\end{aligned}
\qquad ...(5)
$$

Eq. (5) contains the time-varying factor

$$\cos\left(\frac{E_i - E_f}{\hbar}\right)t = \cos 2\pi\nu t\cdot$$

The electron's position therefore oscillates sinusoidally at the frequency

$$\nu = \frac{E_i - E_f}{h} \qquad \text{...(6)}$$

Hence electromagnetic waves are emitted and the frequency of these waves is the same as that of the oscillation. Thus Bohr's postulate $\nu = (E_i - E_f)/h$ follows naturally from the quantum theory of the atom.

13.6 Selection Rules

The integral

$$\int_{-\infty}^{\infty} x\,\psi^*_{of}\,\psi_{oi}\;dx \qquad \text{...(1)}$$

carries all the essential quantum mechanical information about the two atomic states involved in the transition. The intensity of the radiation is proportional to the integral. Hence the general condition necessary for an atom in an excited state to radiate is that the value of the integral is not zero. Transitions for which this integral is finite are called *allowed transitions*. Transitions for which this integral is zero are called *forbidden transitions.*

Consider the transition between the initial and final atomic states of the hydrogen atom, labeled respectively by the quantum numbers (n', l', m'_l) and (n, l, m_l). Let u represent either the x, y or z coordinate. Then the condition for an allowed transition is

$$\int_{-\infty}^{\infty} u\,\psi^*_{n,l,m_l}\,\psi_{n',l',m_l'} \neq 0 \qquad \text{...(2)}$$

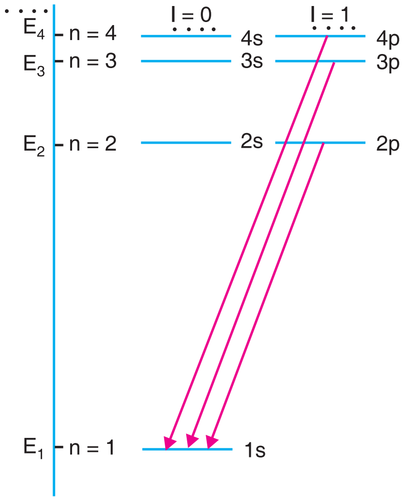


Fig. 13.7

The wave functions $\Psi_{n,\,l,\,m_l}$ for the hydrogen atom are known. Hence Eq. (2) can be evaluated for $u = x$, $u = y$, and $u = z$ for all pairs of states differing in one or more quantum numbers. The selection rules for allowed transitions are

$$\Delta l = \pm 1 \qquad \text{...(3)}$$

$$\Delta m_l = 0, \pm 1 \qquad \text{...(4)}$$

The change in n is not restricted.

Fig. 13.7 shows Lyman transitions allowed by the $\Delta l = \pm 1$ selection rule. All the wavelengths of the Lyman series arise from transitions of the type $np \rightarrow 1s$.

Fig. 13.8 shows that all the wavelengths of the Balmer series are due to transitions connecting $l' = 0$ to $l = 1$ or $l' = 1$ to $l = 0$ or $l' = 2$ to $l = 1$.

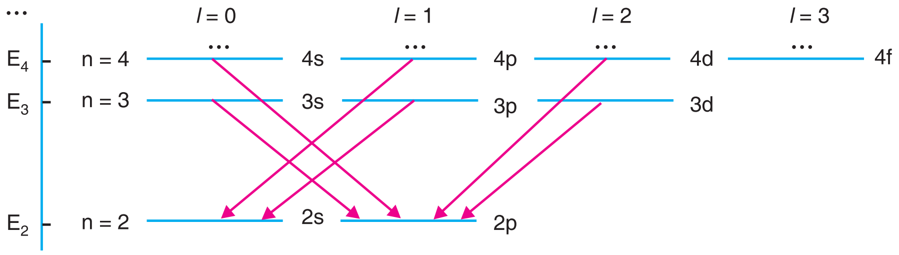
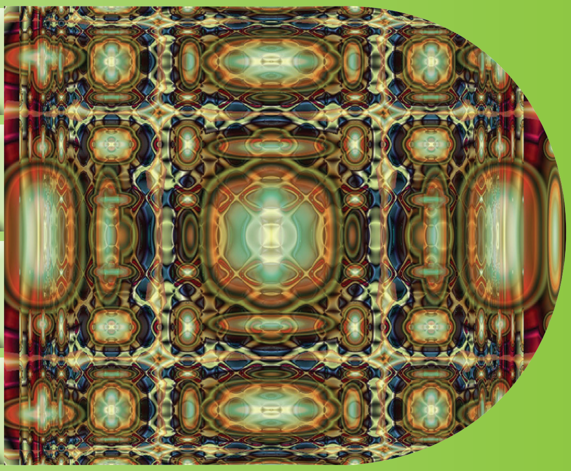


Fig. 13.8

CHAPTER 14



QUANTUM MECHANICS IV

AT A GLANCE

14.1	Finite Square Potential Well	**14.2**	The Parity Operator
14.3	System of Two Identical Particles	**14.4**	Commutator Algebra
14.5	Orbital Angular Momentum Operator		

14.1 Finite Square Potential Well

Fig 14.1 shows a square potential well of width L. Consider a particle with energy E moving along the x-axis. The energy E of the trapped particle is less than the height U of the barriers. Schrodinger's steady-state equation for regions I and III is

$$\frac{d^2\psi}{dx^2} + \frac{2m}{\hbar^2}(E-U)\psi = 0$$

Put $$\frac{\sqrt{2m(U-E)}}{\hbar} = a \qquad \text{...(1)}$$

$$\frac{d^2\psi}{dx^2} - a^2\psi = 0 \quad x < 0, x > L, \qquad \text{...(2)}$$

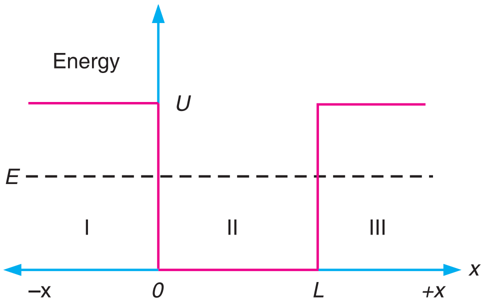


Fig. 14.1

The solutions to Eq. (2) are :

$$\psi_I = A\, e^{ax} + B\, e^{-ax} \qquad \text{...(3)}$$

$$\psi_{III} = C\, e^{ax} + D\, e^{-ax} \qquad \text{...(4)}$$

For ψ to remain finite as $x \rightarrow \pm \infty$, we must have $B = 0$ and $C = 0$.

$$\therefore \qquad \psi_I = A\, e^{ax} \qquad \text{...(5)}$$

$$\psi_{III} = D\, e^{-ax} \qquad \text{...(6)}$$

These wave functions decrease exponentially inside the barriers at the sides of the well.

Schrodinger's equation for region II is

$$\frac{d^2\psi}{dx^2} + \frac{2mE}{\hbar^2}\psi = 0 \qquad \text{...(7)}$$

The solution of Eq. (7) is

$$\psi_{II} = F \sin \frac{\sqrt{2mE}}{\hbar} x + G \cos \frac{\sqrt{2mE}}{\hbar} x \qquad \text{...(8)}$$

At the boundaries $x = 0$ and $x = L$, the wave functions inside and outside must have the same value and the same slope.

$$\psi_{II} = \psi_I \text{ and } \frac{d\psi_{II}}{dx} = \frac{d\psi_I}{dx} \text{ at } x = 0.$$

$$\psi_{II} = \psi_{III} \text{ and } \frac{d\psi_{II}}{dx} = \frac{d\psi_{III}}{dx} \text{ at } x = L.$$

These conditions when applied to Eqs. (5), (6) and (8) yield the following relations :

$$G = A \qquad \text{...(9)}$$

$$F \frac{\sqrt{2mE}}{\hbar} = Aa \qquad \text{...(10)}$$

$$F \sin \frac{\sqrt{2mE}}{\hbar} L + G \cos \frac{\sqrt{2mE}}{\hbar} L = D e^{-aL} \qquad \text{...(11)}$$

$$F \frac{\sqrt{2mE}}{\hbar} \cos\left(\frac{\sqrt{2mE}}{\hbar} L\right) - G \frac{\sqrt{2mE}}{\hbar} \sin\left(\frac{\sqrt{2mE}}{\hbar} L\right) = -Da\, e^{-aL} \qquad \text{...(12)}$$

Eqs. (9) and (10) give

$$F = Ga \frac{\hbar}{\sqrt{2mE}} \qquad \text{...(13)}$$

Eq. (13) together with Eqs. (11) and (12) yields

$$\tan\left(\frac{\sqrt{2mE}}{\hbar} L\right) = \frac{2a(\sqrt{2mE}/\hbar)}{\left(\frac{2mE}{\hbar^2} - a^2\right)}$$

$$\therefore \qquad \tan\left(\frac{\sqrt{2mE}}{\hbar} L\right) = \frac{2\sqrt{E(U-E)}}{2E - U} \qquad \text{...(14)}$$

Only those values of E which satisfy this relation are the allowed energy states. Thus the energy is quantised. The allowed energy values are found by numerical or graphical methods.

The wave functions for the first three allowed energy values and the corresponding probability densities for finding the particle at different locations are shown in Figure 14.2 (*a*) and (*b*) respectively. The particle has a certain probability of being found outside the wall.

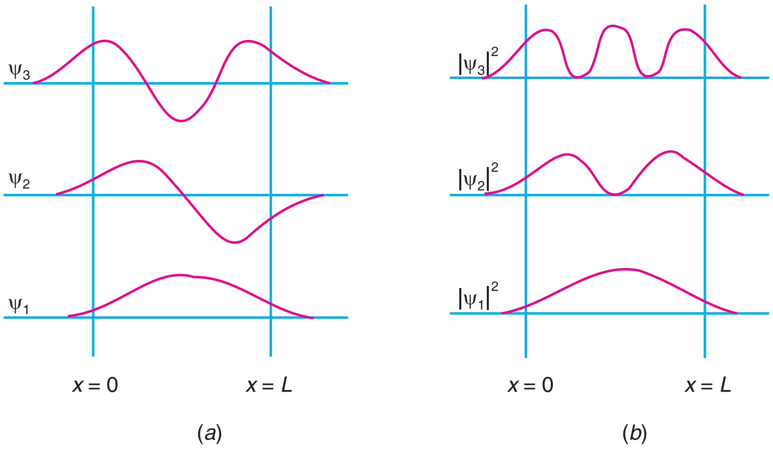


Fig. 14.2

14.2 The Parity Operator

The *Parity* operator P is defined by the relation

$$\mathbf{P}f(\mathbf{r}) = f(-\mathbf{r}) \qquad ...(1)$$

where $f(\mathbf{r})$ is an arbitrary function. Thus the parity operator corresponds to an inversion of the position coordinate $\mathbf{r}$ through the origin.

The parity operation in three dimensions is defined by reflecting all three axes according to the coordinate transformation $(x, y, z) \rightarrow (-x, -y, -z)$ indicated in Figure 14.3. The operation is also called *space inversion* as the transformation changes a right-handed system of coordinates into a left handed system.

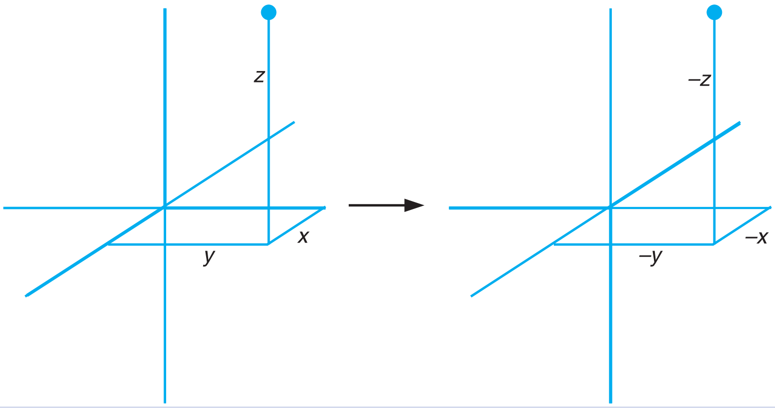


Fig. 14.3

(i) P-operator is Hermitian. For any two wave functions $\phi(\mathbf{r})$ and $\psi(\mathbf{r})$, we have

$$\begin{aligned}\int \phi^*(\mathbf{r}) P\psi(\mathbf{r})\, dr &= \int \phi^*(\mathbf{r})\psi(-\mathbf{r})\, d\mathbf{r} \\ &= \int \phi^*(-\mathbf{r})\psi(\mathbf{r})\, d\mathbf{r} \\ &= \int [P\phi(\mathbf{r})]^*\psi(\mathbf{r})\, d\mathbf{r} \qquad ...(2)\end{aligned}$$

Hence parity operator is a Hermitian operator.

(*ii*) Eigen values of P-operator. Let us now consider the eigenvalue equation for P.

$$P\,\psi_\alpha(\mathbf{r}) = \alpha\,\psi_\alpha(\mathbf{r}) \qquad ...(3)$$

Here α is eigen value of P.

From the definition [Eq. (1)] of P and the fact that $Pf(-\mathbf{r}) = f(\mathbf{r})$, we deduce that

$$P^2 = I \qquad ...(4)$$

where I is the unit operator. Hence

$$P^2\,\psi_\alpha(\mathbf{r}) = \alpha\,P\,\psi_\alpha(\mathbf{r}) = \alpha^2\,\psi_\alpha(\mathbf{r}) = \psi_\alpha(\mathbf{r}) \qquad ...(5)$$

so that $\alpha^2 = 1$ and the eigenvalues of P are $\alpha = \pm 1$.

Denoting the corresponding eigenfunctions by ψ_+ and ψ_-, we have

$$P\,\psi_+(\mathbf{r}) = \psi_+(\mathbf{r}), \qquad P\psi_-(\mathbf{r}) = -\psi_-(\mathbf{r}) \qquad ...(6)$$

or

$$\psi_+(-\mathbf{r}) = \psi_+(\mathbf{r}), \qquad \psi_-(-\mathbf{r}) = -\,\psi_-(\mathbf{r}) \qquad ...(7)$$

Thus $\psi_+(\mathbf{r})$ is an *even* function of **r**, and $\psi_-(\mathbf{r})$ is an odd function of **r**. The eigenfunctions ψ_+ are said to have *even parity,* while the eigenfunctions ψ_- have *odd parity.* We note that ψ_+ and ψ_- are orthogonal.

Any function $\psi(\mathbf{r})$ can always be written as

$$\psi(\mathbf{r}) = \psi_+(\mathbf{r}) + \psi_-(\mathbf{r}) \qquad ...(8)$$

Here,

$$\psi_+(\mathbf{r}) = \frac{1}{2}[\psi(\mathbf{r}) + \psi(-\mathbf{r})] \qquad ...(9)$$

has even parity.

$$\psi_-(\mathbf{r}) = \frac{1}{2}[\psi(\mathbf{r}) - \psi(-\mathbf{r})] \qquad ...(10)$$

has odd parity.

Thus ψ_+ and ψ_- form a complete set.

(*iii*) Commutation relation with Hamiltonian of a system

Under the parity operation $\mathbf{r} \to -\mathbf{r}$, the spherical polar coordinates (r, θ, ϕ) become $(r, \pi - \theta, \pi + \phi)$ [Fig. 14.4].

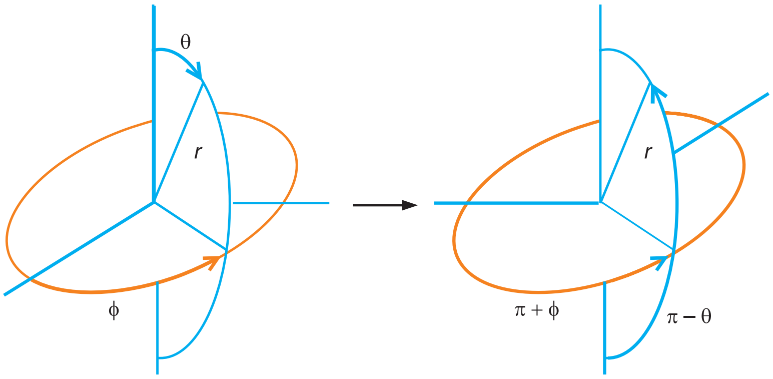


Fig. 14.4

The central force Hamiltonian is

$$H = -\frac{\hbar^2}{2m}\nabla^2 + U(r)$$

$$= -\frac{\hbar^2}{2m}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right] + U(r) \qquad ...(11)$$

It is clearly unaffected by this operation.

In other words, the parity operator P commutes with the Hamiltonian.

$$[P, H] = 0 \qquad ...(12)$$

As a result, simultaneous eigenfunctions of the operators P and H can be found.

Using the expression of $\mathbf{L}^2$, we may write Eq. (11) as

$$H = -\frac{\hbar^2}{2m}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) - \frac{\mathbf{L}^2}{\hbar^2 r^2}\right] + U(r) \qquad ...(13)$$

The Schrodinger equation is

$$\left\{-\frac{\hbar^2}{2m}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) - \frac{\mathbf{L}^2}{\hbar^2 r^2}\right] + U(r)\right\}\psi(r) = E\,\psi(r) \qquad ...(14)$$

Now $\mathbf{L}^2$ and L_z only operate on angular variables, and in addition $[\mathbf{L}^2, L_z] = 0$.

Therefore, we see from Eq. (13) that

$$[H, \mathbf{L}^2] = [H, L_z] = 0 \qquad ...(15)$$

We may thus look for solutions of the Schrodinger equation Eq. (14) which are simultaneous eigenfunctions of the operators H, $\mathbf{L}^2$ and L_z. Since the spherical harmonics $Y_{lm}(\theta, \phi)$ are simultaneous eigenfunctions of $\mathbf{L}^2$ and L_z, we can write a particular solution as

$$\psi_{E,\,l,\,m}(r, \theta, \phi) = R_{E,\,l}(r)\,Y_{lm}(\theta, \phi) \qquad ...(16)$$

Applying the parity operator on the wave function

$$\psi_{E,\,l,\,m}(r, \theta, \phi) = R_{E,\,l}(r)\,Y_{lm}(\theta, \phi), \text{ we have}$$

$$P\,[R_{E,\,l}(r)\,Y_{lm}(\theta, \phi)] = R_{E,\,l}(r)\,Y_{lm}(\pi - \theta, \phi + \pi) \qquad ...(17)$$

Now, from the definition of the spherical harmonics, it can be shown that

$$Y_{lm}(\pi - \theta, \phi + \pi) = (-1)^l\,Y_{lm}(\theta, \phi) \qquad ...(18)$$

so that Y_{lm} has the parity of l. Thus

$$P\,[R_{E,\,l}(r)\,Y_{lm}(\theta, \phi)] = R_{E,\,l}(r)\,(-1)^l\,Y_{lm}(\theta, \phi) \qquad ...(19)$$

The wave function $\psi_{E,\,l,\,m}$ itself has the parity of l (even for even l, odd for odd l).

This symmetry or antisymmetry of a stationary-state wave function under the substitution $\mathbf{r} \to -\mathbf{r}$ is called *the parity of the state.*

14.3 System of Two Identical Particles

Identical particles are those particles in a system for which there will be no change in system by interchanging the particles.

For a system of 2 particles,

$$|\psi|^2(1, 2) = |\psi|^2(2, 1) \qquad ...(1)$$

Wave functions that are unaffected by an exchange of particles are said to be *symmetric.*

$$\psi\,(2, 1) \;=\; \psi\,(1, 2) \qquad ...(2)$$

Wave functions that reverse sign by an exchange of particles are said to be *antisymmetric.*

$$\psi\,(2, 1) \;=\; -\,\psi\,(1, 2) \qquad ...(3)$$

Let particle 1 be in state *a* and particle 2 be in state *b*. Then, the wave function of the system is,

$$\psi_I = \psi_a\,(1)\,\psi_b\,(2) \qquad ...(4)$$

Let particle 2 be in state *a* and particle 1 be in state *b*. Then, the wave function is

$$\psi_{II} = \psi_a\,(2)\,\psi_b\,(1) \qquad ...(5)$$

A linear combination of ψ_{I} and ψ_{II} is the proper description of the system.

The symmetric normalized wave function can be written as

$$\psi_s \;=\; \frac{1}{\sqrt{2}}[\psi_a\,(1)\,\psi_b\,(2) + \psi_a\,(2)\,\psi_b\,(1)] \qquad ...(6)$$

If the wave function of the particles is symmetric, then the particles are said to obey the *Bose Einstein statistics.* Particles of 0 or integral spin have wave functions that are symmetric to an exchange of any pair of them. These particles, which include photons, alpha particles, and helium atoms, do not obey the exclusion principle. This type of particles are known as *bosons.*

The antisymmetric normalized wave function of the system is

$$\psi_A \;=\; \frac{1}{\sqrt{2}}[\psi_a\,(1)\,\psi_b\,(2) - \psi_a\,(2)\,\psi_b\,(1)] \qquad ...(7)$$

Now, the particles are said to obey the *Fermi-Dirac statistics.* They are called *fermions.* Protons, neutrons and electrons are fermions. They have odd half-integral spin. They obey the Pauli exclusion principle.

In Eq. (7) put $a = b$. Then,

$$\psi_A \;=\; \frac{1}{\sqrt{2}}[\psi_a\,(1)\,\psi_a\,(2) - \psi_a\,(2)\,\psi_a\,(1)] = 0 \qquad ...(8)$$

Hence the two particles *cannot* be in the same quantum state. Pauli found that no two electrons in an atom can be in the same quantum state.

14.4 Commutator Algebra

We list here some elementary rules for the calculation of commutators.

If *A*, *B*, and *C* are three linear operators

$$[A, B] \;=\; -\,[B, A] \qquad ...(1)$$

$$[A, B + C] \;=\; [A, B] + [A, C] \qquad ...(2)$$

$$[A, BC] \;=\; [A, B]\,C + B\,[A, C] \qquad ...(3)$$

$$[AB, C] \;=\; [A, C]\,B + A\,[B, C] \qquad ...(4)$$

$$[A, [B, C]] \;+\; [B, [C, A]] + [C, [A, B]] = 0 \qquad ...(5)$$

Now, let us simplify $[A, [B, C]]$.

$$\begin{aligned} [A, [B, C]] &= A\,[B, C] - [B, C]\,A \\ &= A\,(BC - CB) - (BC - CB)\,A \\ &= ABC - ACB - BCA + CBA. \end{aligned}$$

(*a*) Commutation relation between position and momentum

(*i*) Let *x* and P_x be represented as

$$x \to p_{ox}, p_x \to p_{op} = \frac{\hbar}{i}\frac{\partial}{\partial x}.$$

Consider a function $\psi\,(x)$. Then

$$p_x\,\psi = \frac{\hbar}{i}\frac{\partial \psi}{\partial x}$$

and

$$xp_x\,\psi = x\frac{\hbar}{i}\frac{\partial \psi}{\partial x}. \qquad ...(1)$$

Also

$$p_x\,x\psi = \frac{\hbar}{i}\frac{\partial}{\partial x}(x\psi) = \frac{\hbar}{i}\psi + x\frac{\hbar}{i}\frac{\partial \psi}{\partial x} \qquad ...(2)$$

Eq. (1) – Eq. (2) gives,

$$xp_x\,\psi - p_x\,x\psi = x\frac{\hbar}{i}\frac{\partial \psi}{\partial x} - \frac{\hbar}{i}\psi - x\frac{\hbar}{i}\frac{\partial \psi}{\partial x} = -\frac{\hbar}{i}\psi = i\hbar\psi$$

$$\therefore \qquad [x, p_x] = i\,\hbar \qquad ...(3)$$

Now, $\qquad p_y = -i\,\hbar\,\partial/\partial\,y$ and $p_z = -i\,\hbar\,\partial/\partial\,z$.

We have proved the relation $[x, p_x] = i\,\hbar$. More generally, we have

$$[x, p_x] = [y, p_y] = [z, p_z] = i\,\hbar. \qquad ...(4)$$

while all other commutators such as $[x, p_y]$ – vanish. It follows that, for example, x and p_y have common eigenfunctions and can be measured simultaneously with arbitrary accuracy. In contrast, since x and p_x do not commute with each other, a precise simultaneous measurement of both of these observables is impossible.

(*ii*) $[x^2, p_x] = [xx, p_x] = [x, p_x]\,x + x[x, p_x] = i\,\hbar x + xi\,\hbar = 2i\,\hbar\,x$.

(*b*) Commutation relation between Hamiltonian H and momentum p

The Hamiltonian operator for free particle is

$$H_{op} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} \text{ since potential energy } U = 0$$

and

$$p_{op} = \frac{\hbar}{i}\frac{d}{dx}.$$

Consider a function $\psi\,(x)$. Then,

$$p\,\psi = \frac{\hbar}{i}\frac{d\psi}{dx}$$

$$\therefore \qquad Hp\,\psi = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\left(\frac{\hbar}{i}\frac{d\psi}{dx}\right).$$

Similarly,

$$pH\,\psi = \frac{\hbar}{i}\frac{d}{dx}\left[-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2}\right]$$

Solving, $[H, p] = Hp - pH = 0$,

i.e., momentum of a free particle commutes with the Hamiltonian operator.

(*c*) The commutation rules for the components of orbital angular momentum (*L*)

Let L_x, L_y, L_z be cartesian components of $\boldsymbol{L}$.

$$\left.\begin{aligned} L_x &= yp_z - zp_y = y\left(\frac{\hbar}{i}\frac{\partial}{\partial z}\right) - z\left(\frac{\hbar}{i}\frac{\partial}{\partial y}\right) \\ L_y &= zp_x - xp_z = z\left(\frac{\hbar}{i}\frac{\partial}{\partial x}\right) - x\left(\frac{\hbar}{i}\frac{\partial}{\partial z}\right) \\ L_z &= xp_y - yp_x = x\left(\frac{\hbar}{i}\frac{\partial}{\partial y}\right) - y\left(\frac{\hbar}{i}\frac{\partial}{\partial x}\right) \end{aligned}\right\} \qquad \text{...(1)}$$

Now, $[L_x, L_y] = L_x L_y - L_y L_x.$

Substituting operator values

$$\begin{aligned} L_x L_y &= (-i\hbar)^2\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right)\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right) \\ &= -\hbar^2\left\{y\frac{\partial}{\partial z}\left(z\frac{\partial}{\partial x}\right) - y\frac{\partial}{\partial z}\left(x\frac{\partial}{\partial z}\right) - z\frac{\partial}{\partial y}\left(z\frac{\partial}{\partial x}\right) + z\frac{\partial}{\partial y}\left(x\frac{\partial}{\partial z}\right)\right\} \\ &= -\hbar^2\left\{y\frac{\partial}{\partial x} + yz\frac{\partial^2}{\partial z\partial x} - yx\frac{\partial^2}{\partial z^2} - z^2\frac{\partial^2}{\partial y\partial x} + zx\frac{\partial^2}{\partial y\partial z}\right\} \end{aligned}$$

Similarly,

$$\begin{aligned} L_y L_x &= -\hbar^2\left\{\left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}\right)\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right)\right\} \\ &= -\hbar^2\left\{z\frac{\partial}{\partial x}\left(y\frac{\partial}{\partial z}\right) - z\frac{\partial}{\partial x}\left(z\frac{\partial}{\partial y}\right) - x\frac{\partial}{\partial z}\left(y\frac{\partial}{\partial z}\right) + x\frac{\partial}{\partial z}\left(z\frac{\partial}{\partial y}\right)\right\} \\ &= -\hbar^2\left\{zy\frac{\partial^2}{\partial x\partial z} - z^2\frac{\partial^2}{\partial x\partial y} - xy\frac{\partial^2}{\partial z^2} + x\frac{\partial}{\partial y} + xz\frac{\partial^2}{\partial z\partial y}\right\} \end{aligned}$$

But $\dfrac{\partial^2}{\partial x\partial z} = \dfrac{\partial^2}{\partial z\partial x}$ and $\dfrac{\partial^2}{\partial y\partial z} = = \dfrac{\partial^2}{\partial z\partial y}$ and so on.

$$\therefore \quad L_x L_y - L_y L_x = [L_x, L_y] = (-\hbar)^2\left\{y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}\right\}$$

Thus

$$\begin{aligned} [L_x, L_y] &= -i\hbar\left\{-i\hbar\left(y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y}\right)\right\} \\ &= i\hbar\left\{\frac{\hbar}{i}\left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)\right\} \\ &= i\hbar\, L_z \end{aligned} \qquad \text{...(2)}$$

$\hat{L}_x, \hat{L}_y, \hat{L}_z$ do not commute with one another.

$$[\hat{L}_x, \hat{L}_y] = i\hbar\,\hat{L}_z\,;\ [\hat{L}_y, \hat{L}_z] = i\hbar\,\hat{L}_x\,;\ [\hat{L}_z, \hat{L}_x] = i\hbar\,\hat{L}_y$$

(*d*) Commutation relation of L^2 with components L_x, L_y and L_z

We have, $L^2 = L_x^2 + L_y^2 + L_z^2$

$$\begin{aligned}
[L^2, L_x] &= [L_x^2 + L_y^2 + L_z^2, L_x] \\
&= (L_x^2 + L_y^2 + L_z^2)\, L_x - L_x\, (L_x^2 + L_y^2 + L_z^2) \\
&= L_x^2 L_x + L_y^2\, L_x + L_z^2 L_x - (L_x\, L_x^2 + L_x L_y^2 + L_x L_z^2) \\
&= (L_y^2 L_x - L_x\, L_y^2) + (L_z^2 L_x - L_x L_z^2) \\
&= [L_y^2, L_x] + [L_z^2, L_x]
\end{aligned}$$

We know that

$$[AB, C] = [A, C]\, B + A\, [B, C]$$

So,

$$\begin{aligned}
[L^2, L_x] &= [L_y\, L_y, L_x] + [L_z\, L_z, L_x] \\
&= [L_y, L_x]\, L_y + L_y\, [L_y, L_x] + [L_z, L_x]\, L_z + L_z\, [L_z, L_x]
\end{aligned}$$

Now,

$$[L_x, L_y] = i\hbar\, L_z;\ [L_y, L_x] = -i\hbar\, L_z;\ [L_z, L_x] = i\hbar\, Ly.$$

$$\therefore \quad [L^2, L_x] = (-i\hbar\, L_z)\, L_y + L_y\, (-i\hbar\, L_z) + (i\hbar\, L_y)\, L_z + L_z\, (i\hbar\, L_y) = 0.$$

$$\therefore \quad [L^2, L_x] = 0 \qquad \ldots(1\ a)$$

Similarly,

$$[L^2, L_y] = 0 \qquad \ldots(1\ b)$$

$$[L^2, L_z] = 0 \qquad \ldots(1\ c)$$

Hence L^2 commutes with any of the three components of the angular momentum operator.

(*e*) Ladder operators L_+ and L_- : Let us define the operators

$$L_+ = L_x + iL_y \text{ and } L_- = L_x - iL_y \qquad \ldots(2)$$

Commutation Relations of L_z with L_+ and L_-

$$\begin{aligned}
[L_z, L_+] &= [L_z, L_x + iL_y] = [L_z, L_x] + i\, [L_z, L_y] \\
&= i\hbar\ \ L_y + i\, (-i\hbar\, L_x) \\
&= i\hbar\, L_y + \hbar\, L_x \\
&= \hbar\, [L_x + iL_y] = \hbar\, L_+ \qquad \ldots(3)
\end{aligned}$$

and

$$\begin{aligned}
[L_z, L_-] &= [L_z, L_x - iL_y] \\
&= [L_z, L_x] - i\, [L_z, L_y] \\
&= i\hbar\, L_y - i\, (-\, i\hbar\, L_x) \\
&= i\hbar\, L_y - \hbar\, L_x = -\, \hbar\, (L_x - i\, L_y) \\
&= -\, \hbar\, L_- \qquad \ldots(4)
\end{aligned}$$

$$\therefore \quad [L_z, L_\pm] = \pm\, \hbar\, L_\pm \qquad \ldots(5)$$

Commutation Relation of L_+ and L_- mutually

$$\begin{aligned}
[L_-, L_-] &= [(L_x - iL_y), (L_x - iL_y)] \\
&= [L_x, L_x] - i\, [L_x, L_y] + i\, [[L_y, L_x] + [L_y, L_y] \\
&= 0 - i\, (i\hbar\, L_z) + i\, (-\, i\hbar\, L_z\,) + 0 \\
&= \hbar\, L_z + \hbar\, L_z = 2\, \hbar\, L_z. \qquad \ldots(6)
\end{aligned}$$

(*f*) Commutation relation of orbital angular momentum with position

$$\begin{aligned}
[L_x, x] &= [L_x x - xL_x] \\
&= \left\{ y\left(\frac{\hbar}{i}\frac{\partial}{\partial z}\right) - z\left(\frac{\hbar}{i}\frac{\partial}{\partial y}\right) \right\} x - x \left\{ y\left(\frac{\hbar}{i}\frac{\partial}{\partial z}\right) - z\left(\frac{\hbar}{i}\frac{\partial}{\partial y}\right) \right\}
\end{aligned}$$

If ψ (x) is a function of x, we have

$$(L_x\, x - xL_x)\, \psi = \left\{ y\left(\frac{\hbar}{i}\frac{\partial}{\partial z}(x\psi)\right) - z\left(\frac{\hbar}{i}\frac{\partial}{\partial y}(x\psi)\right) \right\} - x \left\{ y\frac{\hbar}{i}\frac{\partial \psi}{\partial z} - z\left(\frac{\hbar}{i}\frac{\partial \psi}{\partial y}\right) \right\}$$

$$= 0$$

or
$$[L_x, x] = 0$$

Similarly,

$$[L_x, y] = i\hbar\, z; \qquad [L_x, z] = -\, i\hbar\, y;$$
$$[L_y, y] = 0; \qquad [L_y, z] = i\hbar\, x;\ [L_y, x] = -\, i\hbar\, z, \qquad \text{...(7)}$$
$$[L_z, y] = -\, i\hbar\, x; \qquad [L_z, z] = 0;\ \ [L_z, x] = i\hbar\, y$$

Similarly, relations between **L** and **p** vectors by keeping in mind the cyclic order of the indices are

$$x \rightarrow y,\ y \rightarrow z,\ z \rightarrow x.$$

Thus
$$[L_x, p_x] = 0,\ [L_x, p_y] = i\hbar\, p_z,\ [L_x, p_z] = -\, i\hbar\, p_y \text{ etc.}$$

14.5 Orbital Angular Momentum Operator

The classical orbital angular momentum of a particle is

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} \qquad \text{...(1)}$$

where **r** and **p** are the position and momentum vectors of the particle, respectively. **p** is represented by the vector operator $-i\hbar\nabla$. So **L** is represented by the vector operator $-\, i\hbar\, (\mathbf{r} \times \nabla)$.

$$\mathbf{L} = \mathbf{r} \times (-\, i\hbar\, \nabla) = -\, i\hbar\, \mathbf{r} \times \nabla \qquad \text{...(2)}$$

Let L_x, L_y, L_z be cartesian components of **L.**

$$\mathbf{i}\, L_x + \mathbf{j}\, L_y + \mathbf{k}\, L_z = -i\hbar \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x & y & z \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \end{vmatrix}$$

$$\therefore \qquad \hat{L}_x = -i\hbar \left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y} \right)$$

$$\hat{L}_y = -i\hbar \left(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z} \right)$$

$$\hat{L}_z = -i\hbar \left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x} \right) \qquad \text{...(3)}$$

Let us now represent Cartesian components L in terms of spherical polar coordinates. The Cartesian coordinates (x, y, z) and spherical polar coordinates (r, θ, ϕ), are related by Fig. 14.5.

$$x = r \sin\theta \cos\phi,$$
$$y = r \sin\theta \sin\phi,$$
$$z = r \cos\theta. \qquad \text{...(1)}$$

These equations yield

$$r^2 = x^2 + y^2 + z^2,\ \cos\theta = \frac{z}{r}.$$

$$\tan\phi = \frac{y}{x} \qquad \text{...(2)}$$

Using Eqs. (1) and (2), we obtain

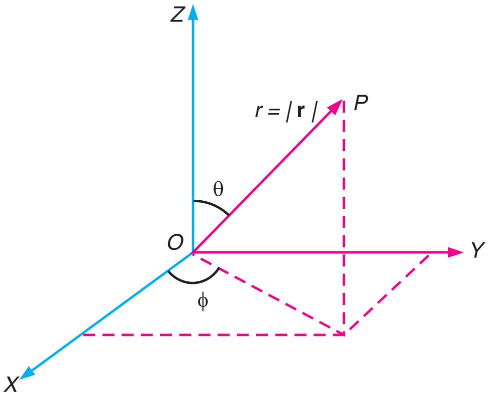


Fig. 14.5

$$\begin{aligned}
\frac{\partial r}{\partial x} &= \sin\theta\cos\phi,\ \frac{\partial r}{\partial y} = \sin\theta\,\sin\phi,\ \frac{\partial r}{\partial z} = \cos\theta \\
\frac{\partial \theta}{\partial x} &= \frac{1}{r}\cos\theta\cos\phi,\ \frac{\partial \theta}{\partial y} = \frac{1}{r}\cos\theta\sin\phi,\ \frac{\partial \theta}{\partial z} = -\frac{1}{r}\sin\theta \qquad \text{...(3)} \\
\frac{\partial \phi}{\partial x} &= -\frac{\sin\phi}{r\sin\theta},\ \frac{\partial \phi}{\partial y} = \frac{\cos\phi}{r\sin\theta},\ \frac{\partial \phi}{\partial z} = 0.
\end{aligned}$$

Using these relations, we obtain

$$\begin{aligned}
\frac{\partial f}{\partial x} &= \frac{\partial f}{\partial r}\frac{\partial r}{\partial x} + \frac{\partial f}{\partial \theta}\frac{\partial \theta}{\partial x} + \frac{\partial f}{\partial \phi}\frac{\partial \phi}{\partial x} \\
&= \sin\theta\cos\phi\frac{\partial f}{\partial r} + \frac{1}{r}\cos\theta\cos\phi\frac{\partial f}{\partial \theta} - \left(-\frac{1}{r}\frac{\sin\phi}{\sin\theta}\frac{\partial f}{\partial \phi}\right) \qquad \text{...(4)} \\
\frac{\partial f}{\partial y} &= \frac{\partial f}{\partial r}\frac{\partial r}{\partial y} + \frac{\partial f}{\partial \theta}\frac{\partial \theta}{\partial y} + \frac{\partial f}{\partial \phi}\frac{\partial \phi}{\partial y} \\
&= \sin\theta\sin\phi\frac{\partial f}{\partial r} + \frac{1}{r}\cos\theta\sin\phi\frac{\partial f}{\partial \theta} + \frac{1}{r}\frac{\cos\phi}{\sin\theta}\frac{\partial f}{\partial \phi}. \qquad \text{...(5)} \\
\frac{\partial f}{\partial z} &= \frac{\partial f}{\partial r}\frac{\partial r}{\partial z} + \frac{\partial f}{\partial \theta}\frac{\partial \theta}{\partial z} + \frac{\partial f}{\partial \phi}\frac{\partial \phi}{\partial z} = \cos\theta\frac{\partial f}{\partial r} - \frac{1}{r}\sin\theta\frac{\partial f}{\partial \theta}. \qquad \text{...(6)} \\
L_x &= i\hbar\left(\sin\phi\frac{\partial}{\partial \theta} + \cot\theta\,\cos\phi\frac{\partial}{\partial \phi}\right) \qquad \text{...(7)} \\
L_y &= i\hbar\left(-\cos\phi\frac{\partial}{\partial \theta} + \cot\theta\,\sin\phi\frac{\partial}{\partial \phi}\right) \qquad \text{...(8)} \\
L_z &= -i\hbar\frac{\partial}{\partial \phi}. \qquad \text{...(9)}
\end{aligned}$$

L^2 in spherical polar coordinates is given by

$$\begin{aligned}
L^2 &= L_x^2 + L_y^2 + L_z^2 \\
&= -\hbar^2\Bigg[\Bigg\{\sin^2\phi\frac{\partial^2}{\partial \theta^2} + \cot^2\theta\cos^2\phi\frac{\partial^2}{\partial \phi^2} + \sin\phi\frac{\partial}{\partial \theta}\left(\cot\theta\cos\phi\frac{\partial}{\partial \phi}\right) \\
&\quad + \cot\theta\cos\phi\frac{\partial}{\partial \phi}\left(\sin\phi\frac{\partial}{\partial \theta}\right)\Bigg\} + \Bigg\{\cos^2\phi\frac{\partial^2}{\partial \theta^2} + \cot^2\theta\sin^2\phi\frac{\partial^2}{\partial \phi^2} \\
&\quad - \cos\phi\frac{\partial}{\partial \theta}\left(\cot\theta\,\sin\phi\frac{\partial}{\partial \phi}\right) + \cot\theta\,\sin\phi\frac{\partial}{\partial \phi}\left(-\cos\phi\frac{\partial}{\partial \theta}\right)\Bigg\} + \left\{\frac{\partial^2}{\partial \phi^2}\right\}\Bigg] \\
&= -\hbar^2\Bigg[\frac{\partial^2}{\partial \theta^2} + (\cot^2\theta + 1)\frac{\partial^2}{\partial \phi^2} + \sin\phi(-\operatorname{cosec}^2\theta)\cos\phi\frac{\partial}{\partial \phi} \\
&\quad + \cot\theta\,\cos^2\phi\frac{\partial}{\partial \theta} - \cos\phi(-\operatorname{cosec}^2\theta)\,\sin\phi\frac{\partial}{\partial \phi} + \cot\theta\,\sin^2\phi\frac{\partial}{\partial \theta}\Bigg] \\
&= -\hbar^2\left[\frac{\partial^2}{\partial \theta^2} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial \phi^2} + \cot\theta\frac{\partial}{\partial \theta}\right]
\end{aligned}$$

$$= -\hbar^2 \left[\frac{\cos\theta}{\sin\theta} \frac{\partial}{\partial\theta} + \frac{\partial^2}{\partial\theta^2} + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial\phi^2} \right]$$

$$\therefore \qquad \mathbf{L}^2 = -\hbar^2 \left[\frac{1}{\sin\theta} \frac{\partial}{\partial\theta} \left(\sin\theta \frac{\partial}{\partial\theta} \right) + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial\phi^2} \right] \qquad ...(10)$$

This is the expression for the operator of the square of the angular momentum.

Expressions for the Eigen values of the square of the total angular momentum (L^2) and its z-component (L_z).

The angular momentum of atoms and molecules is almost as important as the energy as it helps to give physical significance to our mathematical functions. For a one-electron atom, the angular momentum **L** is quantized both in magnitude and direction, with relations

$$|\mathbf{L}|^2 = l(l+1)\hbar^2$$

and

$$L_z = m_l \hbar.$$

Here, $|\mathbf{L}|^2$ is the square of the magnitude of the orbital angular momentum about the centre of the atom and L_z is the component of the angular momentum along the z-axis.

l and m_l are quantum numbers which restrict $|\mathbf{L}|^2$ and L_z to certain specific values.

To prove the above relations for an atom we must show that the wave functions of the atom are eigenfunctions of the angular momentum operators $\hat{\mathbf{L}}^2$ and $\hat{\mathbf{L}}_z$ with eigenvalues $l(l+1)\hbar^2$ and $m_l \hbar$ respectively.

Let us now first apply the operator $\hat{L}_z$ to the one-electron atom wave function

$$\psi(r, \theta, \phi) = R(r)\,\Theta(\theta)\,\Phi(\phi)$$

$$\hat{L}_z \psi = -i\hbar \frac{\partial\psi}{\partial\phi} \quad \text{or} \quad L_z \psi = -i\hbar R\,\Theta \frac{d\Phi}{d\phi}$$

The function $\Phi(\phi)$ for the atom is given by

$$\Phi = A e^{i m_l \phi}$$

$$\frac{d\Phi}{d\phi} = A i m_l e^{i m_l \phi} = i m_l \Phi.$$

Substituting it in the above expression for $\hat{L}_z \psi$, we get

$$\hat{L}_z \psi = -i\hbar\, i m_l R\,\Theta\,\Phi$$

or

$$\hat{L}_z \psi = m_l \hbar \psi$$

This result means that the wave functions ψ of the one-electron atom are the eigen functions of $\hat{L}_z$ having eigen values given by

$$L_z = m_l \hbar.$$

This is the expression for the quantized values of the *z*-component of the angular momentum of the atom. (L_x and L_y, however, do *not* obey quantization relations).

We next apply the operator $\hat{\mathbf{L}}^2$ to the wave function $\psi = R\,\Theta\,\Phi$.

$$\hat{L}^2 \psi = -\hbar^2 \left[\frac{1}{\sin\theta} \frac{\partial}{\partial\theta} \left(\sin\theta \frac{\partial}{\partial\theta} \right) + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial\phi^2} \right] R\,\Theta\,\Phi$$

$$= -\hbar^2 R\left[\frac{\Phi}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{d\Theta}{d\theta}\right)+\frac{\Theta}{\sin^2\theta}\frac{d^2\Phi}{d\phi^2}\right].$$

The Φ-equation is $\Phi = Ae^{im_l\Phi}$

$$\frac{d^2\Phi}{d\phi^2} = A(im_l)^2 e^{im_l\Phi} = -m_l^2\Phi.$$

$$\therefore \qquad \hat{L}^2\psi = -\hbar^2 R\Phi\left[\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{d\Theta}{d\theta}\right)-\frac{m_l{}^2}{\sin^2\theta}\Theta\right].$$

The function Θ (θ) is a solution of the equation

$$\frac{1}{\sin\theta}\frac{d}{d\theta}\left(\sin\theta\frac{d\Theta}{d\theta}\right)+\left[l(l+1)-\frac{m_l^2}{\sin^2\theta}\right]\Theta = 0.$$

This means that the quantity in brackets in the above expression for $\hat{L}^2\psi$ is equal to $-l(l+1)\Theta$. Hence we get

$$\hat{L}^2\psi = -\hbar^2 R\Phi[-l(l+1)\Theta]$$

or

$$\hat{L}^2\psi = l(l+1)\hbar^2 R\Theta\Phi$$

or

$$\hat{L}^2\psi = l(l+1)\hbar^2\psi$$

It means that the wave functions ψ of the one-electron atom are the eigenfunctions of $\hat{L}^2$ having eigenvalues given by

$$|\mathbf{L}|^2 = l(l+1)\hbar^2$$

This is the expression for the quantized values of the square of the angular momentum of the atom.

EXERCISE

1. Discuss quantum mechanically the problem of a particle in a finite square potential well. Draw diagrams showing the amplitude wave and probability density for the same. What will be the effect of increasing the width of the square well on energy levels?
2. Find the value of the following commutations

   (*a*) $[L_x, L_y]$ (*b*) $[L_z, x]$ (*c*) $[L_x, p_y]$ (*d*) $[L^2, L_z]$ (*e*) $[L_z, p_z]$
3. Obtain the commutation relation for L_x, L_y and L_z, the components of angular momentum operator. Show that L^2 commutes with any of the three components.
4. Define a commutator and describe the algebra of commutators.
5. Write operators corresponding to the various rectangular components of angular momentum in spherical polar coordinates and show that

$$\hat{L}^2 = -\frac{h^2}{4\pi^2}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right].$$

   Hence deduce expressions for the eigenvalues of the square of the total angular momentum and its *z*-component.
6. The wave function of a particle in a classically forbidden region is

   (*a*) a sine function (*b*) a cosine function (*c*) a positive exponential (*d*) a negative exponential

Ans. (*d*)

7. Show that the momentum operator $\frac{\hbar}{i}\cdot\frac{\partial}{\partial x}$ is Hermitian. (*Agra* 1973)

SOL. Momentum operator $\hat{p} = \frac{\hbar}{i}\cdot\frac{\partial}{\partial x}$

Complex conjugate of $\hat{p}, \hat{p}* = -\frac{\hbar}{i}\cdot\frac{\partial}{\partial x}$

If $\hat{p}$ is hermitian operator, then in a given state ψ, its average value should be real.

i.e, $\langle \hat{p} \rangle = \int_{-\infty}^{\infty} \psi * \frac{\hbar}{i}\cdot\frac{\partial \psi}{\partial x} dx$ should be real.

Integrating the equation by parts,

$$\langle \hat{p} \rangle = \frac{\hbar}{i}[\psi * \psi]_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \frac{\hbar}{i}.\frac{\partial \psi *}{\partial x}\psi \, dx = \int_{-\infty}^{\infty} \psi\left(-\frac{\hbar}{i}\cdot\frac{\partial \psi *}{\partial x}\right) dx = \langle p* \rangle$$

Since the expectation value of $\hat{p}$ is equal to the expectation value of its complex conjugate $\hat{p}$ *, it is obvious that $\langle \hat{p} \rangle$ is real. Hence the momentum operator $\hat{p}$ is hermitian.

8. An eigenfunction of the operator d^2/dx^2 is $\psi = e^{2x}$. Find the corresponding eigenvalue.

SOL. The eigenvalue equation is $\mathbf{G}\,\psi_n = G_n\,\psi_n$.

Here, **G** is the operator that corresponds to a certain dynamical variable G and each G_n is a real number.

$$\mathbf{G}\psi = \frac{d^2}{dx^2}(e^{2x}) = 4e^{2x} = 4\psi \qquad (\because e^{2x} = \psi)$$

$\therefore$ $\mathbf{G}\psi = 4\psi$.

The eigenvalue is $G = 4$.

9. Which of the following operators are Hermitian operators?

(*a*) position (r) (*b*) momentum (p) (*c*) Total energy (E) (*d*) All of the above.

Ans. (*d*)

10. Find the eigenvalues of the linear momentum operator.

SOL. The eigenvalue equation of linear momentum is

$$-i\hbar\frac{d\psi}{dx} = a\psi; \; (a = \text{eigenvalue})$$

$$\frac{d\psi}{\psi} = \frac{ia}{\hbar}dx$$

Integrating,

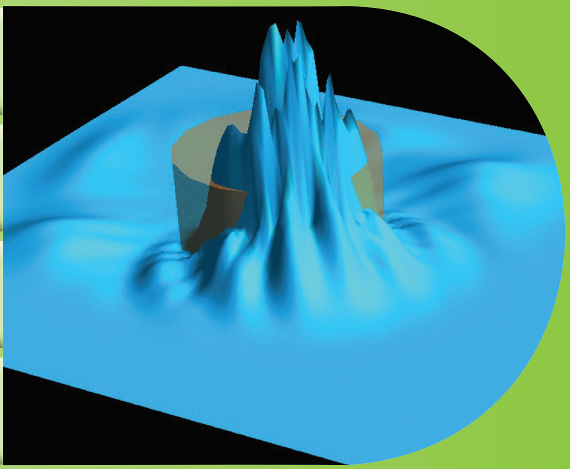
$$\ln \psi = \frac{ia}{\hbar}x + \ln C; \; C = \text{constant}$$

$\therefore$

$$\psi = C\exp\left(\frac{ia}{\hbar}x\right)$$

For ψ to be finite in the region $-\infty$ to $+\infty$, the eigenvalue a has to be real. Hence, all real values of a are the eigenvalues of linear momentum. In other words, linear momentum is not quantized.

CHAPTER
15



QUANTUM MECHANICS V

AT A GLANCE			
15.1	Dirac Delta Function	**15.2**	Properties of δ-Function
15.3	Fourier Transforms	**15.4**	Application of Fourier Transform to a Rectangular Function
15.5	Form of Wave Function in terms of Definite Momentum	**15.6**	Probability Density
15.7	To Show that the Energy Eigen Values are Real	**15.8**	Deuteron Problem
15.9	Dirac's 'BRA' and 'Ket' Vectors: Dirac Notation	**15.10**	Linear Harmonic Oscillator (Operator method)
15.11	Heisenberg's Uncertainty Relation Derived from Operators.	**15.12**	Equation of Motion: Evolution of System with Time; Constants of the Motion
15.13	Angular Momentum	**15.14**	Eigen Values of J^2 and J_z
15.15	Angular Momentum Matrices for $j = \frac{1}{2}$	**15.16**	Angular Momentum Matrices for $j = 1$.
15.17	Addition of Angular Momenta: Clebsch Gordan Coefficients	**15.18**	Addition of Angular Momenta

15.1 Dirac Delta Function

Definition. The Dirac δ-function is defined such that

$$\int_{-\infty}^{\infty} \delta(x-x')\,dx' = 1 \qquad \text{...(1)} \qquad \text{and} \qquad \delta(x-x') = \begin{cases} 0 & (\text{for } x-x' \neq 0) \\ \infty & (\text{for } x-x' = 0) \end{cases} \qquad \text{...(2)}$$

The properties of the δ-function listed in Eqs. (1) and (2) make it clear that the Dirac δ-function is not a function in a rigorous mathematical sense since the integral (if it exists) of a function that is zero everywhere except at one point must vanish.

The *Dirac delta-function* is used in mathematical physics wherever exist functions with non-zero values in very short interval.

EXAMPLE. *Consider the position observable x. The quantum mechanical operator for position is $\hat{x}$. Consider the eigen value equation with this operator $\hat{x}$, operating on the eigen function ψ_0.*

$$\hat{x}\psi_0 = x'\psi_0$$

ψ_0 is the eigen function and it corresponds to a state, in which the *particle is located* precisely at the point $x = x'$. In other words, ψ_0 must be *infinitely concentrated* at $x = x'$ (Fig. 15.1).

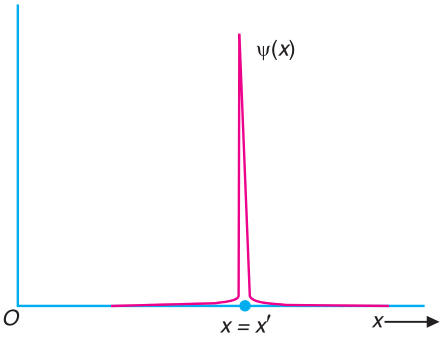


Fig. 15.1

15.2 Properties of δ-Function

1. $\delta(x) = \delta(-x)$
2. $f(x)\,\delta(x-a) = f(a)\,\delta(x-a)$
3. $\int f(x)\delta(x-a) = f(a)$
4. $\delta(ax) = \dfrac{1}{|a|}\,\delta(x)$ for $a > 0$
5. $x\,\delta(x) = 0$

EXAMPLE 1. *Show that the function* $\delta(x) = \lim\limits_{\varepsilon\to 0}\dfrac{\sin(2\pi\varepsilon x)}{\pi x}$ *is a Dirac-delta function.*

Sol. We have
$$\delta(x) = \lim_{\varepsilon\to 0}\frac{\sin(2\pi\varepsilon x)}{\pi x}$$

$$\therefore \quad \delta(x) = 0, \quad \text{when } x \neq 0. \qquad \text{...(1)}$$

and
$$\int_{-\infty}^{\infty}\frac{\sin(2\pi\varepsilon x)}{\pi x}dx = 2\int_{0}^{\infty}\frac{\sin(2\pi\varepsilon x)}{\pi x}dx, \text{ the function being even.}$$

$$= \frac{2}{\pi}\cdot\frac{\pi}{2} = 1 \qquad \text{...(2)}$$

It follows from (1) and (2) that the given function is Dirac-delta function.

15.3 Fourier Transforms

Fourier transforms are very useful in wave packet representation of matter waves. It is often necessary to transform the behaviour of a particle represented by a wave packet from one basis to another basis. For example, the wave packet in configuration space (x) may be required to be expressed in the momentum space (k space). Fourier transform facilitates such transformations from one basis to the other.

The *wave-function integral* is

$$\psi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} g(k)\, e^{ikx}\, dk \qquad ...(1)$$

The *amplitude integral* is

$$g(k) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} \psi(x)\, e^{-ikx}\, dx \qquad ...(2)$$

$\psi(x)$ and $g(k)$, Eqs. (1) and (2), are the Fourier transforms of each other. The function $g(k)$ is the Fourier transform of $\psi(x)$. Similarly, $\psi(x)$ is the Fourier transform of $g(k)$.

A function $\psi(x)$ can only be expressed as a Fourier transform, if the infinite integrals in (1) and (2) *converge.* The integrals converge if

$$\int_{-\infty}^{+\infty} |\psi(x)|^2\, dx < \infty; \text{ and } \int_{-\infty}^{+\infty} |g(k)|^2\, dk < \infty.$$

Jean Baptiste-Joseph Fourier (1768–1830)

15.4 Application of Fourier Transform to a Rectangular Function

A linear combination of harmonic waves of wavelength $\lambda = 2\pi/k$ forms a rectangular function. Here k is called the *propagation constant.*

The amplitude for a rectangular function may be represented as

$$g(k) = \frac{1}{\sqrt{\in}} \text{ for } -\frac{1}{2}\in \le k \le \frac{1}{2}\in$$

$$= 0 \text{ for } |k| > \frac{1}{2}\in \qquad ...(1)$$

The Fourier transform of $g(k)$ is

$$\psi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\in/2}^{+\in/2} g(k)\, e^{ikx}\, dk$$

$$\psi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\in/2}^{\in/2} \frac{1}{\sqrt{\in}}\, e^{ikx}\, dk$$

$$= \sqrt{\frac{2}{\pi\in}}\cdot\frac{\sin(\in x/2)}{x} \qquad ...(2)$$

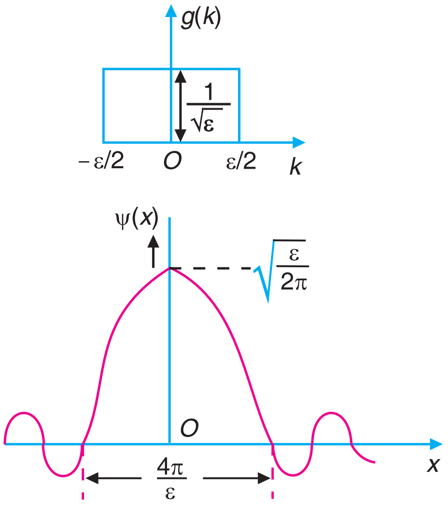


Fig. 15.2

The Fourier transform of $\psi(x)$, *i.e.*, $g(k)$ is given by

$$g(k) = \frac{1}{\sqrt{(2\pi)}}\int_{-\infty}^{+\infty} \sqrt{\left(\frac{2}{\pi\in}\right)}\frac{\sin \in x/2}{x}\, e^{-ikx}\, dx \qquad ...(3)$$

Fig. 15.2 represents rectangular amplitude function and its Fourier transform.

15.5 Form of Wave Function in terms of Definite Momentum

Consider a free particle of total energy E and momentum p moving in the $+x$ direction. The moving particle is associated with de Broglie waves which have wavelength λ and frequency ν.

For de Broglie wave advancing in the positive x direction, the acceptable wave function is

$$\psi(x, t) = A e^{-i(\omega t - kx)} \qquad ...(1)$$

Here A is constant amplitude, ω is constant angular frequency and k is propagation constant.

We have, $\quad E = h\nu = \hbar\omega$ and $p = \hbar k$

$$\therefore \qquad \omega = \frac{E}{\hbar} \text{ and } k = \frac{p}{\hbar}$$

Substituting these values of ω and k in Eq. (1),

$$\psi(x, t) = A e^{-i/\hbar (Et - px)} \qquad ...(2)$$

Eq. (2) describes the wave equivalent of an unrestricted particle of total energy E and momentum p moving in the $+x$ direction.

In three dimensions, Eqs. (1) and (2) become

$$\psi(\mathbf{r}, t) = A e^{-i(\omega t - \mathbf{k.r})}$$

and

$$\psi(\mathbf{r}, t) = A e^{-i/\hbar (Et - \mathbf{p.\ r})}$$

Here, **k** is the propagation vector.

r represents the position vector for the particle.

15.6 Probability Density

Let $\psi(x, t)$ be a normalised wave function of a particle in a given state. $\psi(x, t)$ by itself is sometimes called *Probability amplitude* for the position of the particle.

The product $\psi^*(x, t)\,\psi(x, t)$ is called *probability density* and is real and hence observable. It is the probability of finding the particle in unit interval of space centered about x at time t.

$$\text{Probability density} = \psi^*(x, t)\,\psi(x, t) = |\psi(x, t)|^2 \qquad ...(1)$$

A large value of $|\psi|^2$ means the strong possibility of the particle's presence. A small value of $|\psi|^2$ means the slight possibility of the particle's presence. This interpretation was made by Max Born.

The probability that the particle may be found in the region between x and $(x + dx)$ at time t is

$$P = \psi^*(x, t)\,\psi(x, t)\,dx \qquad ...(2)$$

The probability of finding the particle in a volume element $d\tau = dx\,dy\,dz$ about any point **r** at time t is expressed as

$$P(\mathbf{r}, t)\,d\tau = |\psi(\mathbf{r}, t)|^2\,d\tau. \qquad ...(3)$$

15.7 To Show that the Energy Eigen Values are Real

Consider a particle of mass m moving in one dimension in a potential $V(x)$. The one dimensional Schrodinger wave equation for the state (n) corresponding to the energy eigenvalue E_n is

$$\frac{-\hbar^2}{2m}\frac{d^2\psi_n}{dx^2} + V(x)\psi_n = E_n\psi_n \qquad ...(1)$$

Similarly, for the state (m) corresponding to the energy eigen value E_m,

$$\frac{-\hbar^2}{2m}\frac{d^2\psi_m}{dx^2} + V(x)\psi_m = E_m\psi_m \qquad ...(2)$$

The complex conjugate of Eq. (2) is

$$\frac{-\hbar^2}{2m}\frac{d^2\psi_m^*}{dx^2}+V(x)\psi_m^* = E_m^*\,\psi_m^* \qquad ...(3)$$

Multiplying Eq. (1) by ψ^*_m and Eq. (3) by ψ_n, we have

$$\frac{-\hbar^2}{2m}\psi_m^*\frac{d^2\psi_n}{dx^2}+\psi_m^* V(x)\psi_n = E_n\,\psi_m^*\,\psi_n \qquad ...(4)$$

$$\frac{-\hbar^2}{2m}\psi_n\frac{d^2\psi_m^*}{dx^2}+\psi_n V(x)\psi_m^* = E_m^*\,\psi_m^*\,\psi_n \qquad ...(5)$$

Subtracting Eq. (5) from Eq. (4), we have

$$\frac{-\hbar^2}{2m}\left[\psi_m^*\frac{d^2\psi_n}{dx^2}-\psi_n\frac{d^2\psi_m^*}{dx^2}\right] = (E_n-E_m^*)\,\psi_m^*\,\psi_n$$

$$\frac{-\hbar^2}{2m}\left[\frac{d}{dx}\left(\psi_m^*\frac{d\psi_n}{dx}-\psi_n\frac{d\psi_m^*}{dx}\right)\right] = (E_n-E_m^*)\,\psi_m^*\,\psi_n \qquad ...(6)$$

Integrating Eq. (6) between the limits – ∞ to + ∞, we get

$$\frac{-\hbar^2}{2m}\int_{-\infty}^{+\infty}\frac{d}{dx}\left(\psi_m^*\frac{d\psi_n}{dx}-\psi_n\frac{d\psi_m^*}{dx}\right)dx = (E_n-E_m^*)\int_{-\infty}^{+\infty}\psi_m^*\,\psi_n\,dx \qquad ...(7)$$

$$\frac{-\hbar^2}{2m}\left[\left(\psi_m^*\frac{d\psi_n}{dx}-\psi_n\frac{d\psi_m^*}{dx}\right)\right]_{-\infty}^{+\infty} = (E_n-E_m^*)\int_{-\infty}^{+\infty}\psi_m^*\,\psi_n\,dx$$

Since ψ_m and ψ_n vanish at infinity, the R.H.S. vanishes.

$$\therefore \qquad (E_n-E_m^*)\int_{-\infty}^{\infty}\psi_m^*\,\psi_n\,dx = 0 \qquad ...(8)$$

Put $\quad m = n.\text{Then,}\,(E_n-E_n^*)\displaystyle\int_{-\infty}^{\infty}\psi_n^*\,\psi_n\,dx = 0$

$$(E_n-E_n^*)\int_{-\infty}^{\infty}|\psi_n|^2\,dx = 0$$

The integral is a positive quantity.

$$\therefore \qquad (E_n-E_n^*) = 0$$

$$\therefore \qquad E_n = E_n^*$$

This proves that all energy eigen values are real.

EXAMPLE 1. *Show that any two eigen functions corresponding to different energy eigen values are orthogonal to each other.*

Sol. Let ψ_m and ψ_n correspond to the energies E_m and E_n. Then, from Eq. (8), we have

$$(E_n - E_m^*)\int_{-\infty}^{\infty}\psi_m^*\,\psi_n\,dx = 0$$

If $m \neq n$, $$\int_{-\infty}^{\infty} \psi_m^* \psi_n \, dx = 0$$

This represents the *orthogonality condition.* If two functions ψ_n and ψ_m have this property, they are *orthogonal* to each other.

Example 2. *What are degenerate states?*

Sol. Degeneracy. We have already proved that for every wave function ψ, there must be a corresponding energy value E, *i.e.*, for every eigen function there will be one eigen value. It means that each energy state must have a wave function which will be the characteristic of the system. If for two different ψ's the same value of energy may be obtained, these two states (of different ψ's) will be called *degenerate states.*

The states of a system with different eigen functions, but all corresponding to the same eigen value are known as degenerate states.

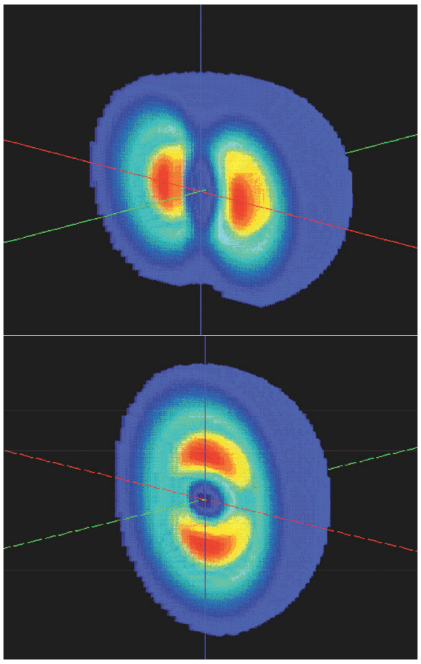
The degree of degeneracy is defined as the number of linearly independent eigen functions corresponding to the same eigen value.

Example 3. *What are orthonormal functions?*

Sol. Orthonormal functions. The functions which are orthogonal and also normalised are called orthonormal functions.

15.8 Deuteron Problem

Deuteron is the nucleus of the deuterium atom (^{2}H or ^{2}D), consisting of one proton and one neutron bound together. Experimentally, the binding energy of the deuteron is found to be 2.22 MeV and the net angular momentum is $J = 1$. This angular momentum is entirely due to the parallel spins of the proton and the neutron. There is no orbital angular momentum, *i.e.*, the system is in an s state.



Graphic Image of a Deuteron.

The reduced mass of the proton-neutron system is

$$\mu = \frac{m_p m_n}{m_p + m_n} = \frac{m_p^{\,2}}{2m_p} = \left(\frac{m_p}{2}\right). \quad (\because \; m_p \approx m_n)$$

The time-independent Schrodinger equation for such a system of zero orbital angular momentum reduces to a simple one-dimensional equation involving the radial variable r,

$$\frac{-\hbar^2}{2\mu}\frac{d^2 u(r)}{dr^2} + U(r)u(r) = Eu(r) \qquad ...(1)$$

For the sake of simplicity, we will assume that the potential is a square well, of infinite height at $r = 0$, and of finite height U_0 at $r = L$ (Fig. 15.3).

We have already studied the solution of the Schrodinger equation for such a square well in Section 14.1. The energy eigen values satisfy the equation,

$$k_1 \cot k_1 L = -k_2 \qquad ...(2)$$

Here, $$k_1 = \sqrt{\frac{2\mu E}{\hbar^2}} \qquad ...(3)$$

$$k_2 = \sqrt{\frac{2\mu(U_0 - E)}{\hbar^2}} \qquad ...(4)$$

$$= \sqrt{\frac{2\mu B}{\hbar^2}} \qquad ...(5)$$

Eq. (2) gives a relation among E, U_0, and L.

The binding energy of the deuteron is known,

$B = 2.22$ MeV.

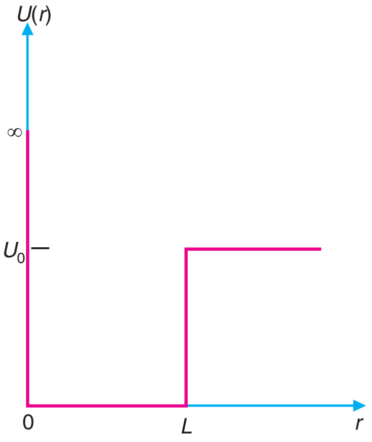
From scattering experiments it is known that

$U_0 = 33$ MeV.

$\therefore\ E = U_0 - B = 30.8$ MeV

$= 30.8 \times 1.6 \times 10^{-13}$ **J** $= 4.93 \times 10^{-12}$ **J**.

Since E and U_0 are known, we can calculate the width of the square well, L, using Eq. (2).



15.3

$$L = \frac{1}{k_1}\cot^{-1}\left(\frac{-k_2}{k_1}\right) = \frac{\hbar}{\sqrt{2\mu E}}\cot^{-1}\left(-\sqrt{\frac{B}{E}}\right) \qquad ...(6)$$

$$= \frac{1\cdot 05 \times 10^{-34}\ \text{J.s}}{\sqrt{2 \times 0\cdot 837 \times 10^{-27}\ \text{kg} \times 4\cdot 93 \times 10^{-12}\ \text{J}}}\cot^{-1}\left(-\sqrt{\frac{2\cdot 22}{30\cdot 8}}\right)$$

$$= 2.13 \times 10^{-15}\ \text{m} \qquad ...(7)$$

This value is in reasonable agreement with the range of the strong potential mentioned above.

A square well with these values of the parameters has only *one* bound state. This is in agreement with the observed absence of excited states in the deuteron.

15.9 Dirac's 'BRA' and 'Ket' Vectors: Dirac Notation

A state of a system is represented by a certain type of vector known *as ket vector* or *ket* represented by the symbol $|\,>$. For example, $|\,a>$ denotes the ket vector corresponding to state a of the system (ψ_a in old notation).

The kets form a linear vector space : any linear combination of several ket vectors is also a ket vector. For example, let us consider two kets $|\,a>$ and $|\,b>$ and two arbitrary complex numbers c_1 and c_2. The linear combination

$$|\,v> = c_1\,|\,a> + c_2\,|\,b> \qquad ...(1)$$

is a vector of *ket space*.

The bra conjugate to ket $|\,a>$ is represented by the symbol $<a\,|$. The bra vector $<\,|$ is analogous to the complex conjugate (or hermitian adjoint) of the wave function of the system. In the Schrodinger representation $|\,\psi\rangle \to \psi(x)$, $\langle\psi| \to \psi^*(x)$.

The bra conjugate to ket $|\,v> = c_1\,|\,a> + c_2\,|\,\mathrm{b}>$ is

$$<v\,| = c_1{}^*<\mathrm{a}\,| + c_2{}^*<\mathrm{b}\,|$$

The scalar product of two state vectors $|\psi_1\rangle$ and $|\psi_2\rangle$ is defined by

$$\langle\psi_1|\psi_2\rangle = \int_{-\infty}^{\infty}\psi_1^*(x)\,\psi_2(x)\,dx \qquad ...(2)$$

If the scalar product of two vectors is zero, the vectors are said to be orthogonal

$$\langle \psi_1 | \psi_2 \rangle = \int_{-\infty}^{\infty} \psi_1^*(x)\,\psi_2(x)\,dx = 0 \qquad \text{...(3)}$$

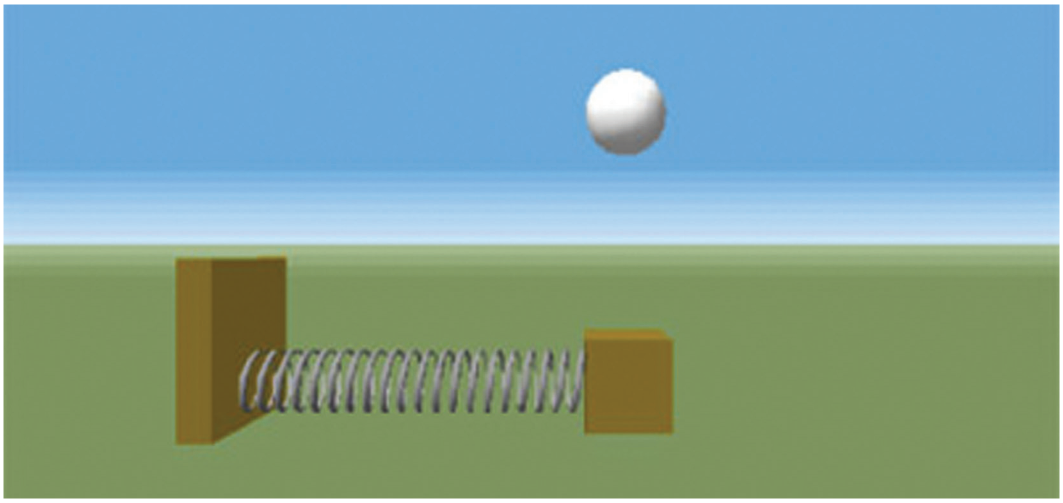
If the scalar product of a vector $|\psi_1\rangle$ with itself is unity, it is said to be normalized.

$$\langle \psi_1 | \psi_1 \rangle = \int_{-\infty}^{\infty} \psi_1^*(x)\,\psi_1(x)\,dx = 1 \qquad \text{...(4)}$$

15.10 Linear Harmonic Oscillator (Operator method)

In Chapter 11 (see Sec. 11.13), we had solved the Schrodinger equation for the linear harmonic oscillator problem. In this section, we will use Dirac's bra and ket algebra to solve the same problem.

For the linear harmonic oscillator problem, the Hamiltonian is given by



Linear Harmonic Oscillator.

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2} m\omega^2 \hat{x}^2 \qquad \text{...(1)}$$

The operators $\hat{x}$ and $\hat{p}$ satisfy the commutation relation :

$$[\hat{x}, \hat{p}] = i\hbar \qquad \text{...(2)}$$

Our objective is to solve the eigenvalue equation

$$\hat{H}|\psi_n\rangle = E_n|\psi_n\rangle \qquad \text{...(3)}$$

We introduce a pair of *matrix operators*

$$\hat{a} = \sqrt{\frac{m\omega^2}{2}}\,.\,\hat{x} + i\sqrt{\frac{1}{2m}}\,\hat{p} \qquad \text{...(4)}$$

and its adjoint.

$$\hat{a}^\dagger = \sqrt{\frac{m\omega^2}{2}}\,.\,\hat{x} - i\sqrt{\frac{1}{2m}}\,\hat{p} \qquad \text{...(5)}$$

Taking the product $\hat{a}^\dagger \hat{a}$, we find that

$$\hat{a}^\dagger \hat{a} = \hat{H} - \tfrac{1}{2}\hbar\omega \qquad \text{...(6)}$$

Similarly
$$\hat{a}\,\hat{a}^{\dagger} = \hat{H} + \tfrac{1}{2}\hbar\omega \qquad \text{...(7)}$$

Using Eqs. (3), (6) and (7), we get :

$$\hat{a}\,\hat{a}^{\dagger}\left|\psi_n\right\rangle = \left(\hat{H} + \tfrac{1}{2}\hbar\omega\right)\left|\psi_n\right\rangle = \left(E_n + \tfrac{1}{2}\hbar\omega\right)\left|\psi_n\right\rangle \qquad \text{...(8)}$$

and
$$\hat{a}^{\dagger}\,\hat{a}\left|\psi_n\right\rangle = \left(\hat{H} - \tfrac{1}{2}\hbar\omega\right)\left|\psi_n\right\rangle = \left(E_n - \tfrac{1}{2}\hbar\omega\right)\left|\psi_n\right\rangle \qquad \text{...(9)}$$

Multiply Eq. (8) by $\hat{a}^{\dagger}$:

$$\hat{a}^{\dagger}\,\hat{a}\,\hat{a}^{\dagger}\left|\psi_n\right\rangle = \left(E_n + \tfrac{1}{2}\hbar\omega\right)\hat{a}^{\dagger}\left|\psi_n\right\rangle \qquad \text{...(10)}$$

This can be written in the form :

$$\hat{a}^{\dagger}\,\hat{a}\left|\psi_{n+1}\right\rangle = \left[(E_n + \hbar\omega) - \tfrac{1}{2}\hbar\omega\right]\left|\psi_{n+1}\right\rangle \qquad \text{...(11)}$$

where

$$\hat{a}^{\dagger}\left|\psi_n\right\rangle = \left|\psi_{n+1}\right\rangle \qquad \text{...(12)}$$

Eq. (11) is just another form of Eq. (9) for $\left|\psi_{n+1}\right\rangle$, provided

$$E_{n+1} = \text{En} + \hbar\,\omega \qquad \text{...(13)}$$

Thus given any eigen vector $\left|\psi_n\right\rangle$, we can generate a new eigen vector $\left|\psi_{n+1}\right\rangle$ through Eq. (12) with the eigen value raised by ħω, in accordance with Eq. (13). This process can be repeated to generate higher and higher eigen values unless E_{n} is the highest level. However, the shape of the harmonic oscillator potential shows that there is no highest level and the procedure for generating higher levels can be continued indefinitely.

Multiplying Eq. (9) by $\hat{a}$, we get :

$$\hat{a}\,\hat{a}^{\dagger}\,\hat{a}\left|\psi_n\right\rangle = \left(E_n - \tfrac{1}{2}\hbar\omega\right)\hat{a}\left|\psi_n\right\rangle \qquad \text{...(14)}$$

This can be written in the form :

$$\hat{a}\,\hat{a}^{\dagger}\left|\psi_{n-1}\right\rangle = \left[(E_n - \hbar\omega) + \tfrac{1}{2}\hbar\omega\right]\left|\psi_{n-1}\right\rangle \qquad \text{...(15)}$$

Here,
$$\hat{a}\left|\psi_n\right\rangle = \left|\psi_{n-1}\right\rangle \qquad \text{...(16)}$$

Eq. (15) is just another form of Eq. (8) for $\left|\psi_{n-1}\right\rangle$, provided

$$E_{n-1} = E_n - \hbar\,\omega \qquad \text{...(17)}$$

Thus given any eigen vector $\left|\psi_n\right\rangle$, we can generate a new eigen vector $\left|\psi_{n-1}\right\rangle$ through Eq. (16) with the eigen value lowered by ħω, in accordance with Eq. (17). This process can be repeated to generate lower and lower eigen vectors unless $\left|\psi_n\right\rangle$ is the ground state $\left|\psi_0\right\rangle$.

For the ground state we have the relation

$$\hat{a}\left|\psi_0\right\rangle = 0 \qquad \text{...(18)}$$

since a further lower level does not exist.

It then follows from Eq. (9) for $n = 0$, that :

$$\left[E_0 - \tfrac{1}{2}\hbar\omega\right]\left|\psi_0\right\rangle = 0 \qquad \text{...(19)}$$

And since $\left|\psi_0\right\rangle \neq 0$, we have

$$\left(E_0 - \tfrac{1}{2}\hbar\,\omega\right) = 0$$

or $$E_0 = \tfrac{1}{2}\hbar\omega \qquad \text{...(20)}$$

This determines the ground state energy.

From Eq. (13), the energy of a general level is then given by :

$$E_n = \left(n+\tfrac{1}{2}\right)\hbar\omega \qquad \text{...(21)}$$

Here, $$n = 0, 1, 2, \ldots$$

$\hat{a}$ and $\hat{a}^{\dagger}$ are the operators which, respectively, lower and raise the energy in the system, in units of $\hbar\omega$. They are therefore called, respectively, the lowering and raising operators.

15.11 Heisenberg's Uncertainty Relation Derived from Operators

Let f and g be two "acceptable" functions. By the Schwarz inequality we have

$$\int f^* f d\tau \int g^* g d\tau \;\geq\; \frac{1}{4}\left[\int (f^* g + g^* f)\, d\tau\right]^2 \qquad \text{...(1)}$$

Now assume a system to be in a state ϕ which need not be an eigen state of any particular operator. We are interested in the results of measurements on the observables belonging to two operators P and Q, at present unspecified.

We introduce the functions

$$f = \left(\hat{P}-\bar{p}\right)\phi \quad \text{and} \quad g = i\left(\hat{Q}-\bar{q}\right)\phi \qquad \text{...(2)}$$

Here, $\bar{p}$ and $\bar{q}$ are mean values associated with $\hat{P}$ and $\hat{Q}$ as

$$\bar{p} = \underset{\tau\to\infty}{Lt} \frac{\int_\tau \phi^* \hat{P} \phi\, d\tau}{\int_\tau \phi^* \phi\, d\tau} \quad \text{and} \quad \bar{q} = \underset{\tau\to\infty}{Lt} \frac{\int_\tau \phi^* \hat{Q} \phi\, d\tau}{\int_\tau \phi^* \phi\, d\tau} \qquad \text{...(3)}$$

Now substituting the values of f and g from Eq. (2) in Eq. (1), we get

$$\int (\hat{P} - \bar{p})^* \varphi^* (\hat{P} - \bar{p})\varphi d\tau \int (\hat{Q} - \bar{q})^* \varphi^* (\hat{Q} - \bar{q})\varphi d\tau$$

$$\geq \frac{1}{4}\left[i\int (\hat{P}-\bar{q})^* \varphi^* (\hat{Q} - \bar{q})\varphi d\tau - i\int (\hat{Q}-\bar{q})^* \varphi^* (\hat{P} - \bar{p})\varphi d\tau\right]^2$$

$$\geq \frac{1}{4}\left[i\int \varphi^* (\hat{P} - \bar{p})(\hat{Q} - \bar{q})\varphi d\tau - i\int \varphi^* (\hat{Q} - \bar{q})(\hat{P} - \bar{p})\varphi d\tau\right]^2 \qquad \text{...(4)}$$

$\hat{P}$ and $\hat{Q}$ are Hermition operators. $\bar{p}$ and $\bar{q}$ are constants. Therefore inequality (4) reduces to

$$\int \varphi^* (\hat{P} - \bar{p})^2 \varphi d\tau \int \varphi^* (\hat{Q} - \bar{q})^2 \varphi d\tau \geq -\frac{1}{4}\left[\int \varphi^* (\hat{P}\hat{Q} - \hat{Q}\hat{P})\varphi d\tau\right]^2 \qquad \text{...(5)}$$

Let us consider the meaning of the quantity

$$\int \varphi^* (\hat{P} - \bar{p})^2 \varphi \tau$$

$$\int \varphi^* (\hat{P} - \bar{p})^2 \varphi d\tau = <(\hat{P} - \bar{p})^2> = <\hat{P}^2> - <2\hat{P}\bar{p}> + <\bar{p}^2>$$

$$= <\hat{P}^2> - 2\bar{p}<\hat{P}> + <\bar{p}^2>$$

$$= <\hat{P}^2> - \bar{p}^2$$

$$= (\Delta p)^2$$

Similarly $\int \varphi^* (\hat{Q} - \bar{q})^2 \varphi \, d\tau = (\Delta q)^2$

Hence $$(\Delta p)^2 (\Delta q)^2 \geq -\frac{1}{4}\left[\int \varphi^* (\hat{P}\hat{Q} - \hat{Q}\hat{P}) \; \varphi \, d\tau\right]^2 \qquad ...(6)$$

If operators $\hat{P}$ and $\hat{Q}$ commute, the right hand side is zero. Since $\hat{P}\hat{Q} = \hat{Q}\hat{P}$ or $\hat{P}\hat{Q} - \hat{Q}\hat{P}$ $= 0$, it is possible for $(\Delta p)^2$ or $(\Delta q)^2$ to be zero or even for both to vanish.

When $\hat{P}$ and $\hat{Q}$ do not commite, relation (6) sets a lower limit for the product of the dispersion, often called uncertaines.

15.12 Equation of Motion: Evolution of System with Time; Constants of the Motion

The state ψ varies with time in a manner determined by the Schrodinger equation

$$\hat{H}\psi = i\hbar \frac{\partial \psi}{\partial t} \qquad ...(1)$$

Here, $\hat{H}$ is the Hamiltonian operator.

Complex conjugate of Eq. (1) is

$$\hat{H}^* \psi^* = -i\hbar \frac{\partial \psi^*}{\partial t} \qquad ... (2)$$

The expectation value of observable associated with an operator $\hat{A}$ in normalized state ψ is

$$< A > = \int \psi^* \hat{A} \psi \, d\tau \qquad ...(3)$$

Taking time derivative of Eq. (3), we have

$$\frac{d}{dt} < A > = \int \frac{\partial \psi^*}{\partial t} \hat{A} \psi \, d\tau + \int \psi^* \hat{A} \frac{\partial \psi}{\partial t} d\tau + \int \psi^* \frac{\partial \hat{A}}{\partial t} \psi \, d\tau \qquad ...(4)$$

From Eq. (1), $$\frac{\partial \psi}{\partial t} = \frac{1}{i\hbar} \hat{H} \psi$$

From Eq. (2), $$\frac{\partial \psi^*}{\partial t} = -\frac{1}{i\hbar} \hat{H}^* \psi^*$$

Using these equations, Eq. (4) becomes

$$\frac{d}{dt} < A > = \int \left(-\frac{1}{i\hbar} \hat{H}^* \psi^* \hat{A} \; \psi \, d\tau\right) + \int \psi^* \hat{A} \; \frac{1}{i\hbar} \hat{H} \psi \, d\tau + \int \psi^* \frac{\partial \hat{A}}{\partial t} \psi \, d\tau$$

$$\frac{d}{dt} < A > = \frac{1}{i\hbar} \int (\psi^* \hat{A}\hat{H}\psi - \hat{H}^* \psi^* \hat{A} \; \psi) d\tau + \int \psi^* \frac{\partial \hat{A}}{\partial t} \psi \, d\tau \qquad ...(5)$$

As $\hat{H}$ is a Hermitian operator, we may write

$$\int H^* \psi^* (A\psi) \; d\tau = \int \psi^* \hat{H}(\hat{A}\psi) d\tau \qquad ...(6)$$

Using Eq. (6), Eq.(5) becomes

$$\frac{d}{dt}<A> \;=\; \frac{1}{i\hbar}\int \psi^{*}(\hat{A}\hat{H}-\hat{H}\hat{A})\,\psi d\tau + \int \psi^{*}\frac{\partial\hat{A}}{\partial t}\psi\, d\tau \qquad ...(7)$$

$$\frac{d}{dt}<A> \;=\; \frac{1}{i\hbar}<(\hat{A}\hat{H}-\hat{H}\hat{A})> \;+ <\frac{\partial\hat{A}}{\partial t}>$$

$$\frac{d}{dt}<A> \;=\; \frac{1}{i\hbar}<[\hat{A},\ \hat{H}]> + <\frac{\partial\hat{A}}{\partial t}> \qquad ...\ (8)$$

Here, $[\hat{A},\hat{H}] = (\hat{A}\hat{H}-\hat{H}\hat{A})$ is called the commutator of $\hat{A}$ and $\hat{H}$.

Thus the rate of change of the *expectation value* of any dynamical variable A may be obtained as the expectation value of

$$(i\hbar)^{-1}[\hat{A},\hat{H}] + \frac{\partial\hat{A}}{\partial t}.$$

If the operator $\hat{A}$ does not depend upon time explicity, then $\frac{\partial\hat{A}}{\partial t}=0$. Eq. (8) reduces to

$$\frac{d}{dt}<A> \;=\; \frac{1}{i\hbar}<[\hat{A},\hat{H}]> \qquad ...(9)$$

Equation (8) or (9) is called *equation of motion in Quantum Mechanics*. This equation is of great importance in observing the integrals of motion of the system.

If commutator of $\hat{A}$ and $\hat{H}$ i.e., $[\hat{A},\hat{H}] = 0$, then from Eq. (9) we have

$$\frac{d}{dt}<A> \;=\; 0 \quad \text{or} \quad <A> = \text{constant}.$$

Thus *if the operator* ($\hat{A}$) *associated with a certain dynamical variable of the system commutes with the Hamiltonian of the system and does not depend upon time explicity, then the dynamical variable is called a constant of motion of the system.*

Whenever it is desired to choose a complete commuting set of observables for any system, it is standard practice to take H as one of them, so that all the observables of the set will be constants of the motion because of their commutability with H. Thus if at some instant of time a system is in a state wherein these observables have specific values, *i.e.*, in a simultaneous eigenstate of the observables, it remains in the same state for all time. The eigenvalues or related quantum numbers which label the state are *good quantum numbers* for the system , since the values of the labels of such a state remain good (unaltered) for all time.

EXAMPLE. *The operator for L_z in polar coordinates is $\hat{L}_z = -i\hbar\frac{\partial}{\partial\phi}$. Therefore, L_z commutes with any operator which does not depend on φ. In particular it commutes with ∇^2 which is independent of φ. Therefore L_z is a constant of the motion for a free particle, whose Hamiltonian is $-\left(\hbar^2/2m\right)\nabla^2$. In fact, it is a constant of the motion even if a potential V is present, provided V is independent of φ, i.e., symmetric with respect to the z axis. Similarly L_x, L_y are constants of the motion if V is symmetric about the x and y axes respectively. If the potential is spherically symmetric, depending on r only and not on θ or φ then all components of the angular mometum are conserved.*

ANGULAR MOMENTUM

15.13 Definition

The total angular momentum is denoted by $\boldsymbol{J} = (\boldsymbol{J_x}, \boldsymbol{J_y}, \boldsymbol{J_z})$ and is defined as the generalized angular momentum operator **J** as any Hermitian operator whose components satisfy the commutation rules,

$$[J_x, J_y] = i\hbar J_z; \; [J_y, J_z] = i\hbar J_x; \; [J_z, J_x] = i\hbar J_y \qquad \text{...(1)}$$

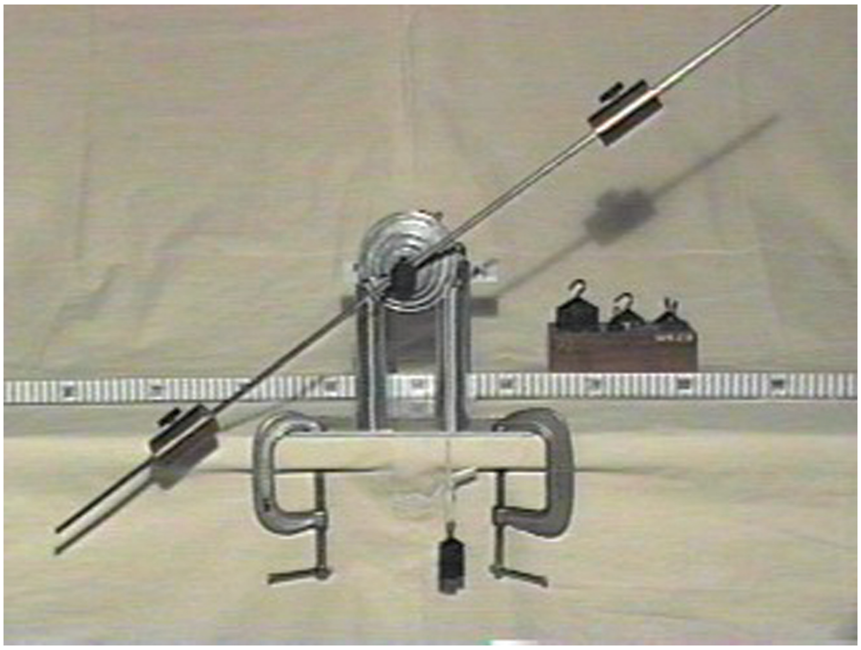
The symbol **J** is introduced here for a *general angular momentum* vector. We use a system of units in which $\hbar = 1$. Thus, Eqs. (1) become

$$[J_x, J_y] = iJ_z; \; [J_y, J_z] = iJ_x; \; [J_z, J_x] = iJ_y \qquad \text{...(2)}$$

As J^2 and J_z commute with each other,

$$[J^2, J_z] = 0 \qquad \text{...(3)}$$

Other equations can be written in cyclic order.



Experimental Setup to Determine Angular Momentum.

15.14 Eigen Values of J^2 and J_z:

Notations for the eigen values of J^2 and J_z. Since J^2 and J_z commute with each other, we can construct a complete set of orthonormal eigenkets which are simultaneous eigen kets of the operators J^2 and J_z belonging to the eigenvalues λ and m respectively. Thus,

$$J^2|\lambda, m\rangle = \lambda|\lambda, m\rangle \qquad \text{...(1)}$$

$$J_z|\lambda, m\rangle = m|\lambda, m\rangle \qquad \text{...(2)}$$

λ and m indicate respectively the eigenvalues of J^2 and J_z and are therefore real numbers.

Ladder operators : We define the *ladder operators*

$$J_+ = J_x + iJ_y \text{ and } J_- = J_x - iJ_y \qquad \text{...(3)}$$

These operators are not hermitian, but they are adjoints of each other :

$$J_+^\dagger = J_- \quad \text{and} \quad J_-^\dagger = J_+ \qquad \ldots(4)$$

These operators satisfy the following commutation relations :

$$[J_z, J_+] = [J_z, J_x] + i\, J_z, J_y = iJ_y + J_x = J_+ \qquad \ldots(5)$$

Similarly, $$[J_z, J_-] = -J_- \qquad \ldots(6)$$

We next calculate the products J_+J_- and J_-J_+.

$$J_+ J_- = J_x^2 + J_y^2 - i\, J_x, J_y = J^2 - J_z^2 + J_z \qquad \ldots(7)$$

Similarly $$J_- J_+ = J^2 - J_z^2 - J_z \qquad \ldots(8)$$

It can be shown that

$$\lambda \geq m^2 \qquad \ldots(9)$$

Effect of the ladder operator on the ket $|\lambda, m\rangle$.

From Eq. (5), $[J_z, J_+] = J_+$

Or $$J_z J_+ - J_+ J_z = J_+ \quad \text{or} \quad J_z J_+ = J_+ J_z + J_+$$

Operating it on the ket $|\lambda, m\rangle$, we get

$$J_z J_+ |\lambda, m\rangle = (J_+ + J_+ J_z)\, |\lambda, m\rangle$$

or $$J_z[J_+ |\lambda, m\rangle] = (m+1)[J_+ |\lambda, m\rangle] \qquad \ldots(10)$$

Further, since J^2 commutes with J_x and J_y we will have

$$J^2, J_\pm = 0$$

or $$J^2 J_\pm |\lambda, m\rangle = J_\pm J^2 |\lambda, m\rangle$$

or $$J^2[J_\pm |\lambda, m\rangle] = \lambda[J_\pm |\lambda, m\rangle] \qquad \ldots(11)$$

Equations (10) and (11) tell us that $J_+ | \lambda, m\rangle$ is a simultaneous eigenket of the operators J^2 and J_z belonging to the eigenvalues λ and $(m + 1)$, provided $J_+ | \lambda, m\rangle$ is not a null ket.

Similarly, using Eq. (6) it can be shown that

$$J_z [J_- |\lambda, m\rangle] = (m - 1) [J_- |\lambda, m\rangle] \qquad \ldots(12)$$

Hence $J_- | \lambda, m\rangle$ is a simultaneous eigenket of J^2 and J_z belonging to the eigenvalues λ and $(m - 1)$ respectively provided $J_- | \lambda, m\rangle$ is not a null ket. Therefore,

$$J_+ |\lambda, m\rangle = C_+ |\lambda, m + 1\rangle \qquad \ldots(13)$$

and $$J_- |\lambda, m] = C_- |\lambda, m - 1\rangle \qquad \ldots(14)$$

Here, C_+ and C_- are normalization constants.

J_+ is a *raising operator* for m.

J_- is a *lowering operator* for m.

So $J_+ |\lambda, m + 1 >$ and $J_-| \lambda, m - 1\rangle$ are proportional to $| \lambda, m + 2\rangle$ and $| \lambda, m - 2\rangle$ respectively.

The repeated application of J_+ must stop at a certain value, say $m = j$, so that the condition $\lambda \geq m^2$ is not violated. Thus

$$J_+ | \lambda, j \rangle = 0 \qquad \ldots(15)$$

Operating on the left by J_-, we obtain [using Eq. (8)]

$$(J^2 - J_z^2 - J_z) | \lambda, j\rangle = 0$$

or $$(\lambda - j^2 - j)|\lambda, j > = 0 \qquad \ldots(16)$$

Since $| \lambda, j\rangle \neq 0$, the number $\lambda - j^2 - j = 0$.

$\therefore$ $$\lambda = j\,(j + 1) \qquad \ldots(17)$$

We can determine the lowest value of m say j', by requiring

$$J_- \,|\, \lambda, j' \,\rangle = 0$$

Operating on the left by J_+ and using Eq. (7), we get

$$(\lambda - j'^2 + j') = 0$$

or $$\lambda = j'\,(j' - 1) \qquad \text{...(18)}$$

Using Eqs. (17) and (18), we get $j'\,(j' - 1) = j\,(j + 1)$.

Thus either $j' = j + 1$ (which we reject since j is supposed to be less then j), or

$$j' = -j \qquad \text{...(19)}$$

Thus, we have a family of kets which correspond to the same eigenvalue $\lambda[= j\,(j + 1)]$ of J^2 but different eigenvalues of J_z. These we denote by

$$|\lambda, j\rangle, |\lambda, j - 1\rangle, \ldots\ |\lambda, -j\rangle \qquad \text{...(20)}$$

Successive application of the ladder operator J_- on $|\lambda, j\rangle$ should ultimately lead to $|\lambda, -j\,\rangle$. (Also successive application of J_+ on $|\lambda, -j\rangle$ should lead to $|\,\lambda,\ j\,\rangle$.) This will happen only if $2j$ is an integer, so that the allowed values of j are

$$j = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots \qquad \text{...(21)}$$

For each value of j the m values go from $-j$ to j in steps of unity. Thus there are $(2j + 1)$ eigenkets of J^2, each having the same eigenvalue $j(j + 1)$. We label the eigenkets as

$$|\,j, m\rangle;\ m = -j, -j + 1, \ldots, j - 1, j \text{ with } j = 0, 1/2, 1, 3/2 \ldots\ldots$$

Any two states $|\,jm\,\rangle$ and $|\,j'm'\rangle$ with $j \neq j'$ or $m \neq m'$ satisfy the orthonormality condition

$$\langle\, j', m'\,|\, j, m > = \delta_{jj'}, \delta_{mm'} \qquad \text{...(22)}$$

It is to be noted that :

(*i*) The only possible values of j are

$$j = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots$$

(*ii*) For a given value of $\boldsymbol{j}$, the only possible values of $\boldsymbol{m}$ are the $(2\boldsymbol{j} + 1)$ numbers :

$$-\boldsymbol{j}, -\boldsymbol{j} + 1, \ldots \boldsymbol{j} - 1, \boldsymbol{j}.$$

15.15 Angular Momentum Matrices for $j = \frac{1}{2}$

Consider the matrix representation of the components of the angular momentum operator (corresponding to $j = \frac{1}{2}$) in a system in which J^2 and J_z are diagonal, *i.e.*, we choose as basis vectors the simultaneous eigenfunctions of J^2 and J_z.

For $$j = \frac{1}{2}, \quad m = \frac{1}{2} \text{ and } -\frac{1}{2}.$$

We have two base states which we denote by $|\,1\,\rangle$ and $|\,2\,\rangle$.

$$|\,1\,\rangle = \left|\frac{1}{2}, \frac{1}{2}\right\rangle = \left|j = \frac{1}{2}, m = \frac{1}{2}\right\rangle \qquad \text{...(1)}$$

and $$|\,2\,\rangle = \left|\frac{1}{2}, -\frac{1}{2}\right\rangle = \left|j = \frac{1}{2}, m = -\frac{1}{2}\right\rangle \qquad \text{...(2)}$$

$$\left.\begin{aligned} J_z\,|1\rangle &= \frac{1}{2}\hbar\,|1\rangle, \quad J_z\,|\,2\,\rangle = -\frac{1}{2}\hbar\,|\,2\,\rangle \\ J^2\,|1\rangle &= \frac{3}{4}\hbar^2\,|\,1\,\rangle, \quad J^2\,|\,2\,\rangle = \frac{3}{4}\hbar^2\,|\,2\,\rangle \end{aligned}\right\} \qquad \text{...(3)}$$

The $(i\,j)^{th}$ matrix element of an operator O is defined by

$$O_{ij} = \langle i \,|\, O \,|\, j \rangle$$

$$\therefore \qquad (J_z)_{12} = \langle 1 \,|\, J_z \,|\, 2 > = 0 = \langle 2 \,|\, J_z \,|\, 1 \rangle = (J_z)_{21}$$

$$(J_z)_{11} = \langle 1 | J_z | 1 \rangle = \frac{1}{2}\hbar, \; (J_z)_{22} = \langle 2 | J_z | 2 \rangle = -\frac{1}{2}\hbar$$

$$J_z = \frac{1}{2}\hbar \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \text{... (4)}$$

Similarly

$$(J^2)_{12} = (J^2)_{21} = 0 \text{ and } (J^2)_{11} = (J^2)_{22} = \frac{3}{4}\hbar^2$$

$$J^2 = \frac{3}{4}\hbar^2 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \text{... (5)}$$

Hence J_z and J^2 matrices contain only principal diagonal terms.

Determination of the matrix elements of J_x and J_y

$$\langle J, m \,|\, J_x \,|\, J, m \rangle = \left\langle j, m \,\middle|\, \frac{1}{2}[J_+ + J_-] \,\middle|\, j, m \right\rangle$$

$$= \frac{1}{2}[(j+m+1)(j-m)]^{\frac{1}{2}}\, \hbar \,\langle j, m' \,|\, j, m+1 \rangle$$

$$+ \frac{1}{2}[(j+m)(j-m+1)]^{\frac{1}{2}}\, \hbar \,\langle j, m' \,|\, j, m-1 \rangle$$

$$= \frac{1}{2}[(j+m+1)(j-m)]^{\frac{1}{2}}\, \hbar\delta_{m',m+1} + \frac{1}{2}[(j+m)(j-m+1)]^{\frac{1}{2}}\, \hbar\delta_{m',m-1}$$

i.e., the matrix element is zero unless $m' = m + 1$ or $m' = m - 1$.

(*i*) $(J_x)_{11} = \langle 1 | J_x | 1 \rangle = \left\langle \frac{1}{2}, \frac{1}{2} \,\middle|\, J_x \,\middle|\, \frac{1}{2}, \frac{1}{2} \right\rangle = 0.$ Similarly $(J_x)_{22} = 0.$

$$(J_x)_{12} = \langle 1 | J_x | 2 \rangle = \left\langle \frac{1}{2}, \frac{1}{2} \,\middle|\, J_x \,\middle|\, \frac{1}{2}, -\frac{1}{2} \right\rangle = \frac{1}{2}\left[\left(\frac{1}{2} - \frac{1}{2} + 1\right)\left(\frac{1}{2} + \frac{1}{2}\right)\right]^{\frac{1}{2}} \hbar = \frac{1}{2}\hbar = (J_x)_{21}$$

$$J_x = \frac{1}{2}\hbar \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \text{... (6)}$$

(*ii*) $(J_y)_{11} = (J_y)_{22} = 0. \quad (J_y)_{12} = -\frac{i}{2}\hbar$ and $(J_y)_{21} = \frac{i}{2}\hbar.$

$$J_y = \frac{1}{2}\hbar \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \text{... (7)}$$

To complete the picture, we should also take note of the kind of entities on which these matrices act. The representative of an arbitrary state $|\psi\rangle$ with respect to the $|\, jm\rangle$ basis is a column whose elements are given by $< jm \,|\, \psi >$.

The matrices of Eqs. (4), (5), (6) and (7) would act to the right on the particular segment of this infinite column, for which $j = \frac{1}{2}$. The matrices can also act *to the left* on the representative of any bra vector $\langle\psi|$, which is a row with elements $\langle \psi \,|\, j'm' \rangle$.

15.16 Angular Momentum Matrices for $j = 1$

For $j = 1, \quad m = 1,\ 0$ and -1.

We have three base states which we denote by $|1\rangle, |2\rangle$ and $|3\rangle$:

$$|1\rangle = |j = 1, m = 1\rangle = |1, 1\rangle$$

$$|2\rangle = |j = 1, m = 0\rangle = |1, 0\rangle$$

$$|3\rangle = |j = 1, m = -1\rangle = |1, -1\rangle$$

(*i*) $J_z\,|1\rangle = \hbar\,|1\rangle\,;\quad J_z\,|2\rangle = 0\,;$ and $J_z\,|3\rangle = -\hbar\,|1\rangle$

$$J_z = \hbar\begin{pmatrix}1 & 0 & 0\\ 0 & 0 & 0\\ 0 & 0 & -1\end{pmatrix} \qquad ...(1)$$

(*ii*) $J_z^2\,|1\rangle = 2\hbar^2\,|1\rangle;\quad J_z^2\,|2\rangle = 2\hbar^2\,|2\rangle\,;$ and $J_z^2\,|3\rangle = 2\hbar^2\,|3\rangle$.

$$J^2 = 2\hbar^2\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix} \qquad ...(2)$$

The matrices representing J^2 and J_z are diagonal.

Similarly, $J_x = \dfrac{\hbar}{\sqrt{2}}\begin{pmatrix}0 & 1 & 0\\ 1 & 0 & 1\\ 0 & 1 & 0\end{pmatrix}$ and $J_y = \dfrac{\hbar}{\sqrt{2}}\begin{pmatrix}0 & -i & 0\\ i & 0 & -i\\ 0 & i & 0\end{pmatrix}$

The procedure is similar to that in the previous section.

The matrix elements are calculated using the fact that $J_x = \dfrac{1}{2}(J_+ + J_-)$ and $J_y = -\dfrac{1}{2}i(J_+ - J_-)$.

ADDITION OF ANGULAR MOMENTA : CLEBSCH GORDAN COEFFICIENTS

15.17 Introduction

We often have to deal with systems in which the total angular momentum is composed of two or more parts that are independent of each other, e.g., the particles with spin (in the non-relativistic limit), systems containing two or more particles such as many electron atoms, etc. The problem is to find how the total angular momentum is related to its component parts.

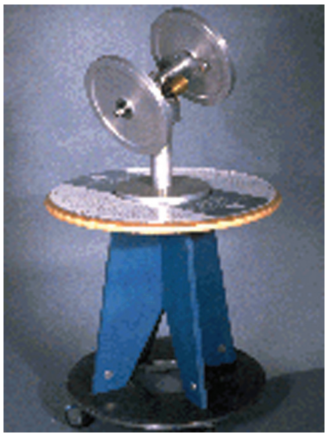
Let there be two non-interacting systems (1) and (2) with angular momenta $\boldsymbol{J_1}$ and $\boldsymbol{J_2}$ and state vectors Ψ_1 and Ψ_2. If J represents the sum of J_1 and J_2 then

$$\boldsymbol{J} = \boldsymbol{J_1} + \boldsymbol{J_2} \qquad ...(1)$$

If the systems (1) and (2) are considered together as a single system, the state vector of combination is represented as

$$\psi = \psi_1\,\psi_2 \qquad ...(2)$$

Given the angular momenta j_1, j_2 of the individual systems, what are the possible values of j for the composite system ?



Dumbbell Gyroscope allows the Addition of Angular Momentum because of its two Wheels.

The answer is stated in terms of the *Vector model* :

If $\boldsymbol{J_1}, \boldsymbol{J_2}$ are thought of as two ordinary vectors of length j_1 and j_2 respectively, these vectors together with $\boldsymbol{J_1} + \boldsymbol{J_2}$ form a triangle (Fig. 15.4). The values of j_1, j_2, j must then be such that a triangle with these sides can be constructed. This *triangle rule* or *triangle condition* immediately sets the maximum and minimum limits for j as $j_1 + j_2$ and $|j_1 - j_2|$ (corresponding to $\boldsymbol{J}_1$ and $\boldsymbol{J}_2$ being parallel and antiparallel respectively). Since angular momentum is quantized, the only allowed values between these limits are $j_1 + j_2 - 1$, $j_1 + j_2 - 2, \ldots, |j_1 - j_2| + 1$.

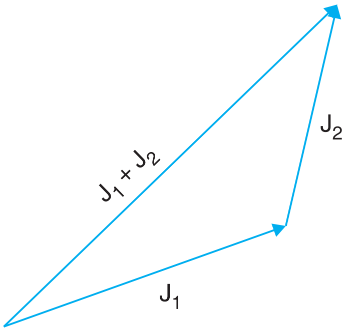


Fig. 15.4

15.18 Addition of Angular Momenta

Let $\boldsymbol{J}_1$ and $\boldsymbol{J}_2$ be two angular momentum operators satisfying the commutation relations :

$$[J_{1x}, J_{1y}] = iJ_{1z} \qquad \ldots(1)$$
$$[J_{2x}, J_{2y}] = iJ_{2z}$$

Other equations can be written in cyclic order.

We use a system of units in which $\hbar = 1$.

We also have $[J_1^2, J_{1x}] = 0 = [J_2^2, J_{2x}]$, etc.

We assume that the components of $\boldsymbol{J}_1$ and $\boldsymbol{J}_2$ commute with each other, *i.e.*,

$$[J_{1x}, J_{2x}] = 0 = [J_{1x}, J_{2y}], \text{ etc.}$$

J is the *total angular momentum operator* for the *composite* system made up of $\boldsymbol{J}_1$ and $\boldsymbol{J}_2$. We define

$$\boldsymbol{J} = \boldsymbol{J}_1 + \boldsymbol{J}_2 \qquad \ldots(2)$$

The components of $\boldsymbol{J}$ satisfy the same commutation relations as that of an angular momentum operator :

$$[J_y, J_z] = iJ_x \qquad \ldots(3)$$
$$[J^2, J_z] = 0, \text{ etc.} \qquad \ldots(4)$$

Here $$J^2 = J_x^2 + J_y^2 + J_z^2 = (J_{1x} + J_{2x})^2 + \ldots + \ldots = J_1^2 + J_2^2 + 2\boldsymbol{J}_1 \cdot \boldsymbol{J}_2 \qquad \ldots(5)$$

Since the total angular momentum vector is a constant of motion, we construct vectors which are simultaneous eigenvectors of J^2 and J_z.

Let $\psi_1 (j_1, m_1)$ be the simultaneous eigenvectors of J_1^2 and J_{1z} i.e.,

$$J_1^2 \psi_1 (j_1, m_1) = j_1 (j_1 + 1) \psi_1 (j_1, m_1) \qquad \ldots(6)$$

and $$j_{1z} \psi_1 (j_1, m_1) = m_1 \psi_1 (j_1, m_1) \qquad \ldots(7)$$

Let $\psi_2(j_2, m_2)$ be the simultaneous eigenvectors of J_2^2 and J_{2z}.

The operators J_1^2, J_2^2, J_{1z} and J_{2z} commute with each other.

The product functions

$$\psi(j_1, j_2, m_1, m_2) = \psi_1(j_1, m_1) \psi_2 (j_2, m_2) \qquad \ldots(8)$$

represent a set of simultaneous eigenvectors of J_1^2, J_2^2, J_{1z} and J_{2z}.

For given values of j_1 and j_2 there will be $(2j_1 + 1)$ possible values of m_1 and $(2j_2 + 1)$ possible values of m_2. Thus there will be $(2j_1 + 1)(2j_2 + 1)$ possible eigenvectors.

The operators J_1^2, J_2^2 J^2 and J_z also commute with each other. Thus we can construct vectors $\phi(j_1, j_2, j, m)$ which are simultaneous eigenvectors of the operators J_1^2, J_2^2, J^2 and J_z belonging to the eigenvalues $j_1 (j_1 + 1), j_2(j_2 + 1), j(j + 1)$ and m respectively.

The allowed values j of the total angular momentum, resulting from the addition of two given angular momenta j_1, j_2, are

$$j_1 + j_2, (j_1 + j_2 - 1), \ldots, |j_1 - j_2| \quad \text{...(9)}$$

For each value of j there will be $(2j + 1)$ values of m. So the total number of ϕ functions (for given values of j_1 and j_2) is

$$\sum_{j=|j_1-j_2|}^{j=j_1+j_2} (2j+1) = (2j_1+1)(2j_2+1) \quad \text{...(10)}$$

The number of ψ functions is also the same.

We can choose either the ψ functions or the ϕ functions as the base states and can represent each of the ψ functions as a linear combination of ϕ functions and conversely. Now the ψ functions as well as the ϕ functions form orthonormal sets. If we normalize them, then the transformation

$$\phi(j_1, j_2, j, m) = \sum_{m_1}\sum_{m_2} C(j_1, j_2; j, m, m_1, m_2)\, \psi(j_1, j_2, m_1, m_2)$$

is unitary. ...(11)

The coefficients

$$C(j_1, j_2; j, m, m_1, m_2) = \langle j_1, j_2, m_1, m_2 \mid j_1, j_2, j, m \rangle$$

are called *Clebsch-Gordan coefficients*.

EXERCISE

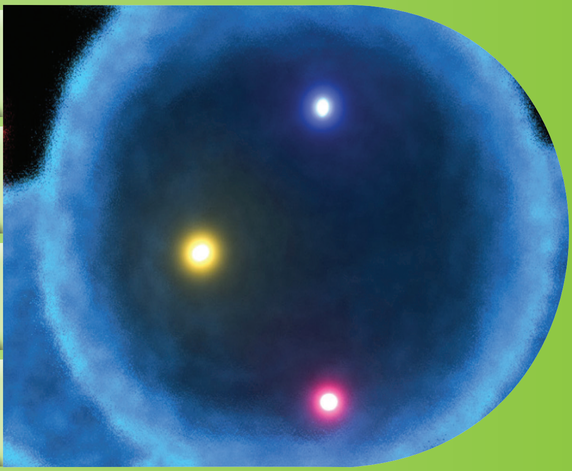
1. Define Dirac delta-function. State the properties of Dirac delta-function.
2. What is Fourier transform? What is its use ?
3. Apply Fourier's transform to obtain the momentum wave-functions for a rectangular function having the amplitude

$$g(k) = \frac{1}{\sqrt{\in}} \text{ for } -\frac{\in}{2} \le k < \frac{\in}{2}$$

$$= 0 \text{ for } |k| > \in / 2.$$

4. Derive an expression for the wave function for a particle with well-defined momentum p and energy E.
5. Define 'probability density'. What is Born's interpretation of wave function?
6. Prove that the energy eigen values are real.
7. Set up Schrodinger equation for deuteron. Discuss the solution of the equation and show how the binding energy, depth of potential well and its width are related.

CHAPTER 16



QUANTUM MECHANICS VI

AT A GLANCE			
16.1	The Helium Atom	**16.2**	Hydrogen Molecule
16.3	Quantum Mechanical Hamiltonian of the Hydrogen Atom	**16.4**	The Harmonic Oscillator
16.5	Comparison of Classical and Quantum Ideas	**16.6**	Scattering Theory
16.7	Definitions of Differential Scattering Cross-section and Total Scattering Cross-section		
16.8	The Scattering Amplitude	**16.9**	Partial Wave Analysis
16.10	Transformation from Centre of Mass to Laboratory Frame of Reference		
16.11	Reduction of the Two-Body Problem into One-Body Problem		

16.1 The Helium Atom

The helium atom consists of a nucleus of charge + 2e and two electrons each of charge – e. Coordinates used in the formulation of the helium Hamiltonian are shown in Fig. 16.1. $\mathbf{r}_1$ and $\mathbf{r}_2$ are the position vectors of the two electrons with respect to the nucleus as origin, and $r_{12} = |\mathbf{r}_1 - \mathbf{r}_2|$ is the distance between the two electrons. m is the electron mass. We shall neglect the effects of spin-orbit interaction and the motion of the nucleus.

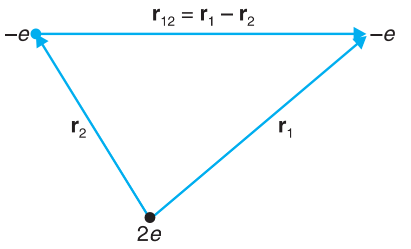
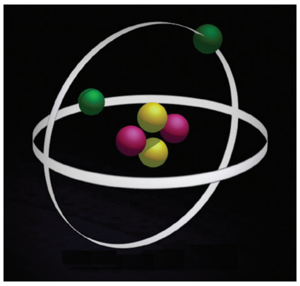


Fig. 16.1



Helium Atom.

The total system Hamiltonian is given by

$$\hat{\boldsymbol{H}} = -\frac{\hbar^2}{2m}(\nabla_1^2+\nabla_2^2)-\frac{2e^2}{4\pi\varepsilon_0 r_1}-\frac{2e^2}{4\pi\varepsilon_0 r_2}+\frac{e^2}{4\pi\varepsilon_0|\mathbf{r}_1-\mathbf{r}_2|} \quad ...(1)$$

The first term is the sum of the two kinetic energy operators.

The second and third terms are the potential energies of the electrons in the field of the nucleus with nuclear charge 2*e*.

The last term represents the Coulomb repulsive energy of the two electrons.

The Schrodinger equation for the states of helium

$$H\psi = E\psi$$

where H is given by Eq. (1) cannot be solved analytically. So we are forced to use approximate methods. The simplest approach is to use first order perturbation theory. We split the Hamiltonian H of Eq. (1) into two parts.

$$H = H_0 + U \quad ...(2)$$

Here, $$H_0 = -\frac{\hbar^2}{2m}(\nabla_1^2+\nabla_2^2)-\frac{2e^2}{4\pi\varepsilon_0}\left(\frac{1}{r_1}+\frac{1}{r_2}\right) \quad ...(3)$$

and $$U = \frac{e^2}{4\pi\varepsilon_0|\mathbf{r}_1-\mathbf{r}_2|} \quad ...(4)$$

H_0 is the Hamiltonian for two non-interacting electrons in the field of the nucleus.

The perturbation U is the electrostatic repulsion between the two electrons. We note that U is not substantially smaller than H_0. Therefore, while we expect our solutions to be qualitatively correct we do not expect very great quantitative accuracy.

As H_0 contains only terms which apply either to one electron or the other, but not to both, it has solutions of the form

$$H_0\,\psi = E\psi \quad ...(5)$$

Here, $$E = E_{n_1 l_1} + E_{n_2 l_2} \quad ...(6)$$

The lowest energy for H_0 occurs when $n_1 = n_2 = 1$ and $l_1 = l_2 = 0$.

The ground state energy is given by

$$E_{0gs} = 2E_{10} \quad ...(7)$$

where $$E_{10} = -4e^2/4\pi\varepsilon_0\,2a_0 \quad ...(8)$$

$$\therefore \quad E_{0gs} = -8\left(\frac{e^2}{4\pi\varepsilon_0\,2a_0}\right) = -8\times 13.6\,\text{eV} = -108.8\,\text{eV}. \quad ...(9)$$

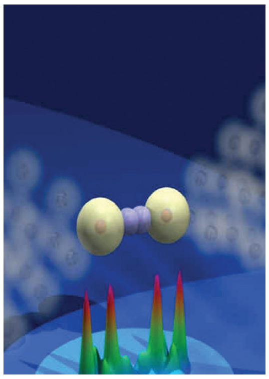
16.2 Hydrogen Molecule

The H_2 *molecule* consists of two hydrogen nuclei and two electrons.

The nuclei (protons) will be labeled A, B and the two electrons “1” and “2” (Fig. 16.2).

r_{12} is the distance between two electrons 1 and 2.

We assume that the nuclei of the two hydrogen atoms A and B are fixed in space a distance $\mathbf{R}_{AB}$ apart and z-axis is chosen parallel to the line joining the two nuclei. $\mathbf{r}_{1A}$ is the position coordinate of the electron 1 from nucleus A. $\mathbf{r}_{2B}$ is the position coordinate of the electron 2 from nucleus B.



Hydrogen Molecule.

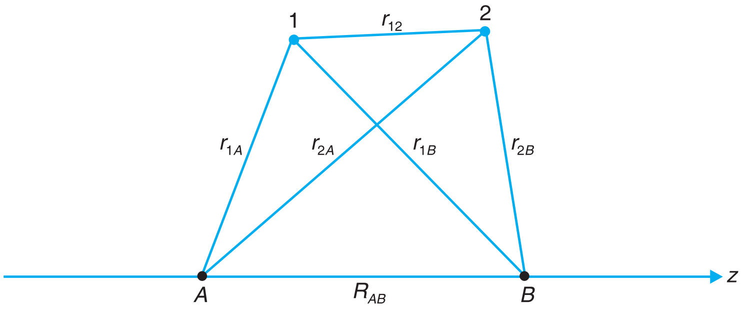


Fig. 16.2

The Hamiltonian for the two electron system is given by

$$H = -\frac{\hbar^2}{2m}\left(\nabla_1^2 + \nabla_2^2\right) + \frac{e^2}{4\pi\varepsilon_0}\left(\frac{1}{R_{AB}} + \frac{1}{r_{12}} - \frac{1}{r_{1A}} - \frac{1}{r_{2B}} - \frac{1}{r_{1B}} - \frac{1}{r_{2A}}\right)$$

16.3 Quantum Mechanical Hamiltonian of the Hydrogen Atom

The quantum-mechanical problem is to be solved by finding the solution of the Schrodinger equation

$$H\psi = E\psi \qquad \text{...(1)}$$

Here, H is the quantum-mechanical Hamiltonian

$$H = -\frac{\hbar^2}{2m}\nabla^2 - \frac{Ze^2}{4\pi\varepsilon_0 r} \qquad \text{...(2)}$$

The term $-[Ze^2/(4\pi\varepsilon_0 r)]$ depends only on the radial distance r. e is the magnitude of the electron charge. m is the electron mass.

We express the operator ∇^2 in Eq. (2) in spherical polar coordinates.

$$\nabla^2 = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{r^2 \sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2}{\partial\phi^2} \qquad \text{...(3)}$$

If we insert Eq. (3) into Eq. (2) we obtain for the quantum-mechanical Hamiltonian

$$H = -\frac{\hbar^2}{2m}\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) - \frac{\hbar^2}{2mr^2}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right] - \frac{Ze^2}{4\pi\varepsilon_0 r} \qquad \text{...(4)}$$

16.4 The Harmonic Oscillator

The traditional ideal harmonic oscillator is shown in Fig. 16.3. The displacement of the mass from its equilibrium position is x, and its momentum is p.

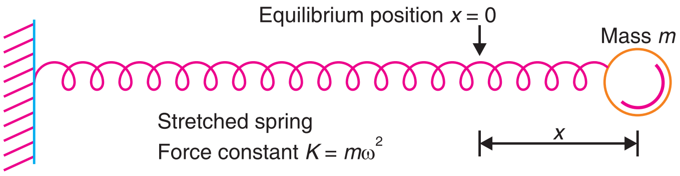


Fig. 16.3

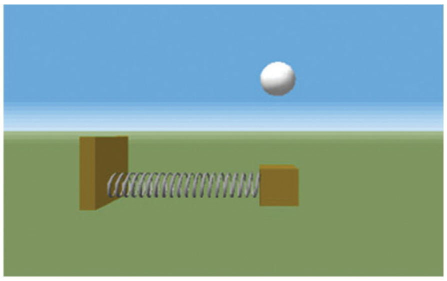
A harmonic oscillator is characterized by a natural oscillation frequency, ω.

$$\omega = \sqrt{K/m} \qquad \text{...(1)}$$

From classical mechanics, the Hamiltonian for such an oscillator is

$$H = \frac{p^2}{2m} + \frac{Kx^2}{2} \qquad \text{...(2)}$$

This algebraic combination of the dynamical variables x and p produces the energy of the system in classical physics. From that fact we expect the quantum mechanical energies of the system to be generated by the operator $\hat{H}$, where



Simple Harmonic Oscillator.

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{K\hat{x}^2}{2} \qquad \text{...(3)}$$

The spectrum of energies we seek is given by the eigenvalue problem for E.

$$\hat{H}|E\rangle = E|E\rangle \qquad \text{...(4)}$$

The eigenvalues (permitted values of the total energy) are given by

$$E = E_n = \left(n + \frac{1}{2}\right)\hbar\omega$$

16.5 Comparison of Classical and Quantum Ideas

1. Wave-particle Duality

The phenomenon of reflection, refraction, interference, diffraction and polarisation of light prove that light possesses '**wave nature**'. On the other hand, photo-electric effect, Compton effect and the emission, absorption of radiation are the phenomena which prove that light possesses quantum or **particle nature**. So light possesses **dual nature**, *i.e.*, it behaves like a wave and a particle.

At atomic levels the determination of *e/m* in case of electrons, protons and ionised atoms provide the experimental tests about the particle or quantum nature. On the other hand the diffraction experiments of Davisson and Germer, G.P. Thomson, double slit interference, prove that electrons, protons, neutrons and other atomic particles possess wave nature. This situation led to think that matter also possesses dual nature. The result of duality of light and matter gave rise to wave-mechanics. This mechanics is widely applicable at atomic and nuclear dimensions. It was independently formulated by Schrodinger and Heisenberg under the names of wave-mechanics and quantum mechanics.

Wave-particle duality is the critical concept on which all of quantum mechanics is built. De Broglie said that to the particle properties of momentum and energy correspond the wave properties of wavelength and frequency.

It is wave-particle duality that destroys our picture of atoms as merely very small beads. Atoms exhibit wave properties. Wave properties of macroscopic beads are never seen! Only for sufficiently small particles does their wave-particle duality become observable.

2. Probability

In describing atomic phenomena we are concerned with probabilities rather than certainties. For example, according to quantum mechanics, the radius of the innermost orbit of the electron in the hydrogen atom is not exactly 5.3×10^{-11} m. On the other hand, it is the most probable radius of the orbit. Therefore, if the radius of the orbit is measured with a suitable device different trials will yield different results, some larger and some smaller than the above value. But the value that is most likely to be found is 5.3×10^{-11} m.

In macroscopic bodies there are very large numbers of individual atoms. Therefore, departure from average behaviour is not noticeable. That is why observed quantities appear to be certainties in Newtonian mechanics. Thus Newtonian mechanics turns out to be an approximation of quantum mechanics. In quantum mechanics we have a single set of principles which apply to both microscopic and macroscopic systems.

3. Quantization of Energy and Angular Momentum

(*i*) The existence of quantum numbers and discrete energy levels was proved for "the infinite square well potential". The discrete levels and quantum numbers are characteristic of *all* problems where a particle is bound to a small region of space. Classical mechanics has no such requirement of discrete energy levels for bound systems. The contrast between quantum and classical behaviour is extreme at lower values of n. For high n, the two treatments converge.

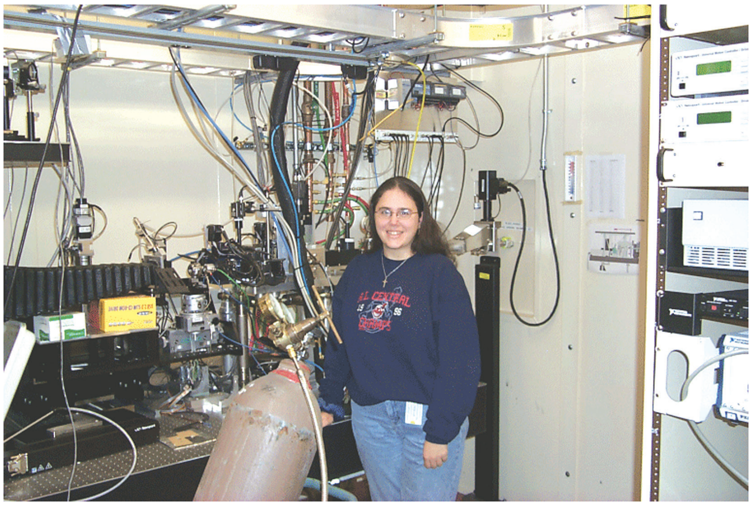
(*ii*) In the hydrogen atom, the boundary conditions, that ψ must be single valued and finite, lead to two important results when applied to the angular part of the wave function. The total angular momentum is quantized as $\left(\sqrt{l(l+1)}\right)\hbar$. The component of angular momentum in the z-direction is quantized as $m_l\,\hbar$, where $m_l \le l$.

SCATTERING THEORY

16.6 Introduction

Scattering theory plays a very important role in quantum mechanics particularly in the field of high energy physics.

Scattering experiments have provided most of the information about interactions among atoms, nuclei and particles. The general technique is to have a collimated beam of projectiles (particles) with well defined energies and sometimes other properties, such as polarization, as well. The particles are then scattered by a target of atoms. The scattered particles are detected by means of a suitable detector. The measurement could be simply the intensity variation as a function of angle or could involve more detailed information like the change in the energy or the polarization state, etc.



Small Angle Scattering Setup Studying the Effect of Pressure on Hollow Nano Co Spheres.

The quantum mechanical interpretation of the scattering process is rather different from that in a classical case.

The motion of a particle cannot be described with complete accuracy by classical orbits. We must use wave packet whose average coordinates give the classical orbits.

Since the particle cannot be localised, a well defined impact parameter cannot be assigned to it.

The scattering must therefore, be described by the wave functions that are solutions of the Schrodinger equations. The probability of scattering is computed by determining the flux at a point. The scattering process is best described by a quantity called the *scattering cross section*. The most commonly used parameter which displays the results of scattering experiments is the *scattering cross-section*. In Sec.16.7 we define the scattering cross-section.

16.7 Definitions of Differential Scattering Cross-section and Total Scattering Cross-section

Consider an incident beam of free particles represented by a plane wave advancing along the *z*-direction (Fig.16.4.).

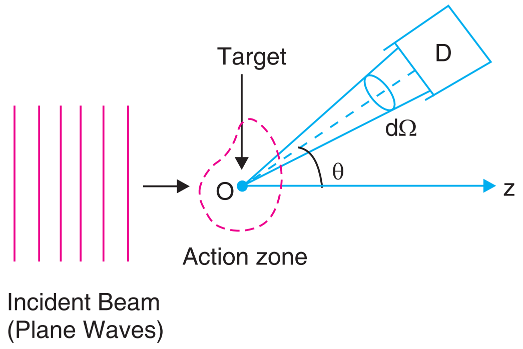

Target
D
dΩ
θ
O
z
Action zone
Incident Beam
(Plane Waves)

Fig. 16.4

The target (scatterer) is supposed to be at the origin. The scattering potential is localised around the origin *O,* producing a small action zone. A detector *D* is placed away from the action zone in a direction defined by the polar angles θ and ϕ.

We assume that the flux is low enough that any interaction of the incident particles with one another is completely negligible, *i.e.*, the scattering of each particle is independent of others.

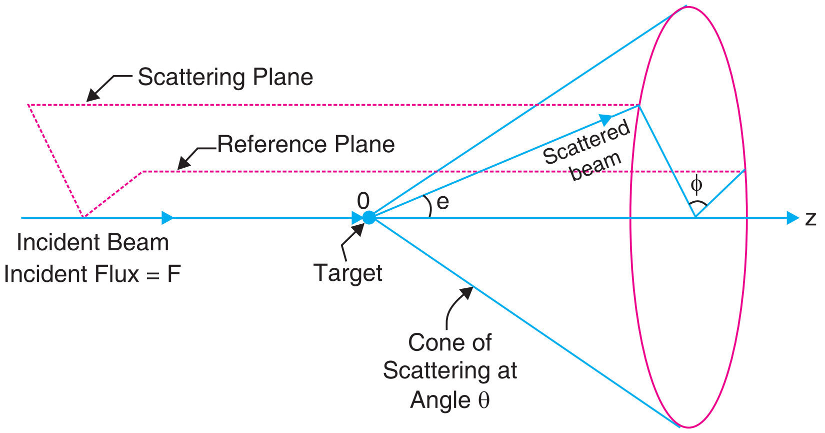

Scattering Plane
Reference Plane
Scattered beam
ϕ
0
e
z
Incident Beam
Incident Flux = F
Target
Cone of
Scattering at
Angle θ

Fig 16.5. Shows a scattering event.

We choose a coordinate system with the origin O at the position of the target or *scattering centre*, and with the z-axis in the direction of the indicent beam.

The direction of any scattered particle is indicated by polar angles (θ, φ), with the z-axis taken as the polar axis. Then θ is the *angle of scattering*, *i.e.*, the angle between the scattered and incident directions. These two directions together define the *plane of scattering*. The azimuthal angle φ specifies the oriention of this plane with respect to some reference plane conaining the z-axis.

Let F denote the incident flux, *i.e.*, the number of particles crossing per second a unit surface placed perpendicular to the beam direction Oz away from O *in the negative* z-direction. A steady incident beam is maintained for an indefinitely long time, *i.e.*, the incident flux F is independent of time. Then there will be a steady stream of scattered particles.

Let ΔN be the number of particles scattered in time Δt into a small solid angle $\Delta\Omega$ about the direction (θ, φ). The number ΔN must be proportional to $\Delta\Omega\,\Delta t$ and to the incident flux F.

We can thus write:

$$\Delta N = \frac{d\sigma(\theta,\varphi)}{d\Omega}.\,\Delta t\,\Delta\Omega.F \qquad \text{...(1)}$$

Here, $(d\sigma/d\Omega)$ is the constant of proportionality. It depends in general on θ and φ. It is called the *differential scattering cross-section*. It depends only on the parameters of the incident particle and nature of the target. The aim of scattering theory is to determine $(d\sigma/d\Omega)$ as a function of θ and φ. $(d\sigma/d\Omega)$ has the dimensions of an area. It is measured in barns and its submultiples :

$$1\ barn = 10^{-28}\ m^2.$$

The total scattering *cross-section* σ is obtained from differential scattering cross-section by integration over all directions:

$$\sigma = \int\left(\frac{d\sigma}{d\Omega}\right)d\Omega = \int_0^{2\pi}\int_0^{\pi}\frac{d\sigma}{d\Omega}.\sin\theta\,d\theta\,d\varphi \qquad \text{...(2)}$$

In most of the cases we consider $(d\sigma/d\Omega)$ is independent of φ.

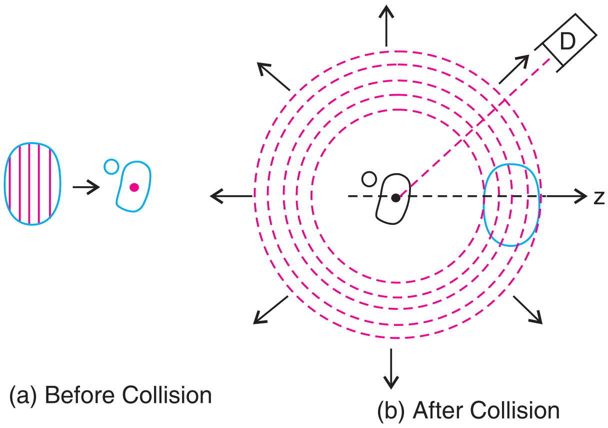
Then Eq. (2) becomes,

$$\sigma = \int\frac{d\sigma}{d\Omega}.2\pi\sin\theta\,d\theta \qquad \text{...(3)}$$

16.8 The Scattering Amplitude

To show that the differential scattering cross-section is equal to the square of the Scattering Amplitude.

When the particles involved in the scattering process are quantum mechanical objects, we must describe them by a wave function. The wave function will be a stationary one if we assume that a steady (time-independent) incident beam is maintained for all time. The phenomenon of scattering is manifested as a distortion in the stationary wave pattern, caused by the presence of the scattering centre.



(a) Before Collision (b) After Collision

Fig. 16.6

As the incident wave packet enters the action zone, the scattering potential *V(r)* becomes effective and the structure of the wave packet is greatly modified. However, once the wave packet comes out of the action zone, it again has a simple form : It consists of a transmitted wave along *Oz* and a scattered wave (Fig. 16.6).

At *large distances* from the scattering centre ($r \to \infty$), the total asymptotic wavefunction $\psi(x)$ is given by :

$$\psi(x) \xrightarrow[r\to\infty]{} \psi_{inc} + \psi_{sc} \qquad \text{... (1)}$$

(*i*) The first term in Eq. (1), ψ_{inc}, represents the parallel beam of incident particles.

(*ii*) The second term, ψ_{sc}, represents the scattered particles moving radially outwards from the centre.

The beam of incident particles with momentum $p = \hbar k$ along the *z*-axis is described by the plane wave (momentum eigenfunction)

$$\psi_{inc} = e^{ikz} \qquad \text{...(2)}$$

Here, *k* is given by

$$\frac{\hbar^2 k^2}{2\,m} = E \qquad \text{...(3)}$$

$|\psi_{inc}|^2$ is the number of incident particles per unit volume.

The incident flux is obtained by multiplying this quantity by the particle velocity *v.*

$$F = |\,\psi_{inc}\,|^2\, v = \hbar k/m \qquad \text{...(4)}$$

We assume that the scattered particles have the same momentum as the incident particles, i.e., the scattering is *elastic*. Then the wave ψ_{sc} representing the scattered particles must have the same propagation constant *k.*

The scattered wave is a *spherical wave* diverging outwards from the origin. It is described by

$$\Psi_{sc} \propto e^{ikr} \qquad \text{...(5)}$$

Further, the flux of scattered particles, $|\,\Psi_{sc}\,|^2\,(\hbar k/m)$ must decrease as $(1/r^2)$ as *r* increases.

$$\Psi_{sc} \propto (1/r) \qquad \text{...(6)}$$

The scattered wavefunction can thus be written in the form :

$$\Psi_{sc} = f\,(\theta, \varphi)\,\frac{e^{ikr}}{r} \qquad \text{...(7)}$$

The function $f(\theta, \varphi)$ is called the *scattering amplitude*.

Let $\Delta \boldsymbol{N}$ be the number of particles scattered in time Δt into a small solid angle $\Delta\Omega$ about the direction (θ, ϕ).

$$\Delta N = (\text{radial flux}) \times \Delta S\,\Delta t \qquad \text{...(8)}$$

Here, $\Delta S \equiv r^2\,\Delta\Omega$ is the element of area (normal to the radial direction) covered by the solid angle $\Delta\Omega$.

$$\therefore \qquad \Delta N = |\,\Psi_{sc}\,|^2\,(\hbar k/m).r^2\,\Delta\Omega\,\Delta t$$

$$\text{or} \qquad \Delta N = |\,f\,(\theta, \varphi\,|^2\,(\hbar k/m).\,\Delta\Omega\,\Delta t \qquad \text{...(9)}$$

From Eq. (1) of Sec. 16.7, we have

$$\Delta \boldsymbol{N} = \frac{d\sigma\,(\theta, \phi)}{d\Omega}\,\Delta \boldsymbol{t}\,\Delta\Omega.\boldsymbol{F} \qquad \text{...(10)}$$

$$\text{or} \qquad \Delta \boldsymbol{N} = \frac{d\sigma\,(\theta, \phi)}{d\Omega}\,\Delta t\,\Delta\Omega.\frac{\mathbf{h}k}{m} \text{ using Eq. (4)} \qquad \text{...(11)}$$

From Eq. (9) and (11), we obtain

$$\frac{d\sigma(\theta, \phi)}{d\Omega} = |f(\theta, \varphi)|^2 \qquad \text{...(12)}$$

The differential scattering cross-section is thus merely the square of the scattering amplitude.

In order to calculate the scattering cross-section, we have to determine the particular stationary wave function $\Psi(x)$ which has the following asymptotic form :

$$\Psi(x) \xrightarrow[r\to\infty]{} e^{ikz} + f(\theta, \varphi)\,\frac{e^{ikr}}{r} \qquad \text{...(13)}$$

Since any stationary wave function must satisfy the time independent Schrodinger equation, we must have

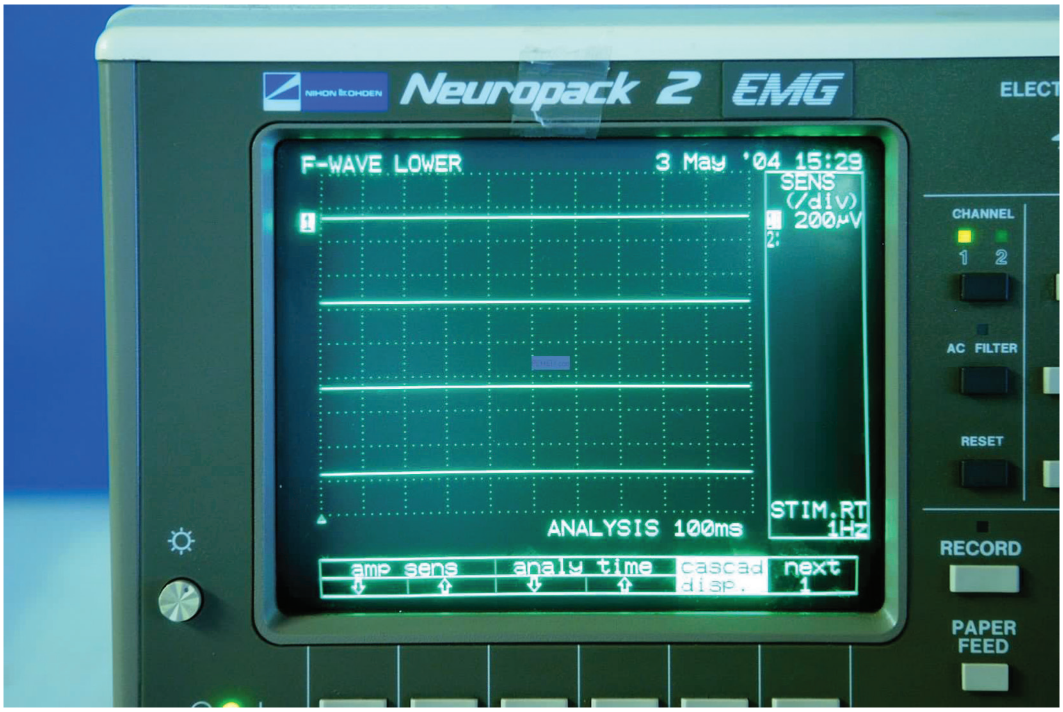
$$-\frac{\hbar^2}{2m}\nabla^2\Psi(x) + V(x)\,\Psi(x) = E\Psi(x), \qquad \text{...(14)}$$

Here, ***m*** represents the mass of the particle. $V(x)$ is the potential energy function for the projectile particle in the force field of the scattering centre.

The eigenvalue E has to be taken equal to the kinetic energy of the particle in the asymptotic region (where $V = 0$). If the solution of Eq. (14) subject to the asymptotic form Eq. (13) can be obtained, then the coefficient of (e^{ikr}/r) in Eq. (13) gives the scattering amplitude.

16.9 Partial Wave Analysis

The method of partial waves is based upon an expansion of the wave function in terms of angular momentum eigenfunctions. It is applicable to spherically symmetric potentials. It is useful only for low-energy incident particles.


NIHON KOHDEN
Neuropack 2 EMG
ELECT
F-WAVE LOWER
3 May '04 15:29
SENS (/div) 200μV
ANALYSIS 100ms
STIM.RT 1Hz
amp sens
analy time
cascad disp.
next 1
CHANNEL
1 2
AC FILTER
RESET
RECORD
PAPER FEED

Wave Analysis Screen.

Consider a plane wave incident along z-axis in a region having interaction potential function $V(r)$. Then total wave function may be expressed as

$$\Psi(r) = \underset{\text{incident wave}}{e^{ikz}} + \underset{\text{scattered wave}}{f(\ )\frac{ikr}{\ }} \qquad \text{...(1)}$$

It is a solution of three dimensional Schroedinger equation

$$\nabla^2 \Psi + \frac{2\mu}{\hbar^2}[E - V(r)]\Psi = 0 \qquad \text{...(2)}$$

The solution of Eq. (2) may be expressed as

$$\Psi(r) + \sum_{l=0}^{\infty} R_l(r)\, Y_{lo}^{(\theta)} \qquad \text{...(3)}$$

This is a superposition of a number of waves. Each term in above equation is called a **partial wave**, corresponding to a particular value of l.

The functions $R_l(r)$ are called radial wave functions.

Put $$\frac{2\mu E}{\hbar^2} = k^2 \quad \text{and} \quad \frac{2\mu V(r)}{\hbar^2} = U(r).$$

Eq. (2) becomes,

$$\nabla^2 \Psi + [k^2 - U(r)]\Psi = 0 \qquad \text{...(4)}$$

As there is symmetry about polar axis *i.e.*, z-axis ($m = 0$) and potential energy function does not involve ϕ, the solution of Eq. (4) may be expressed as

$$\Psi(r, \theta, \phi) = \Psi(r, \theta) \sum_l R_l(r)\, P_l(\cos\theta) \qquad \text{...(5)}$$

Put $$\chi_l = r R_l(r) \qquad \text{...(6)}$$

Eq. (5) becomes,

$$\Psi(r, \theta) = \sum_l r^{-1} \chi_l(r)\, P_l(\cos\theta) \qquad \text{...(7)}$$

Here, $\chi_l(r)$ satisfies the equation

$$\frac{d^2\chi_l}{dr^2} + \left[k^2 - U(r) - \frac{l(l+1)}{r^2}\right]\chi_l = 0 \qquad \text{...(8)}$$

In the asymptotic region, U and l terms in Eq. (8) may be neglected in comparison with k^2. So Eq. (8) becomes,

$$\frac{d^2\chi_l}{dr^2} + k^2\chi_l = 0. \qquad \text{...(9)}$$

or $$\chi_l = e^{\pm ikr} \qquad \text{...(10)}$$

which gives radial wave only.

For better approximation, we define a distance "a" such that when $r > a$, $V(r) = 0$ and when $r < a$, $V(r)$ is appreciably finite. Then Eq. (8) becomes

$$\frac{d^2\chi_l}{dr^2} + \left[k^2 - \frac{l\,(l+1)}{r^2}\right]\chi_l = 0 \qquad \text{...(11)}$$

This is spherical Bessel equation.

On solving Eq. (11), we get

$$R_l(r) = \frac{A_l}{kr}\sin\left(kr - \frac{l\pi}{2} + \delta_l\right) \qquad \text{...(12)}$$

Here, $\boldsymbol{A_l}$ is a constant.

δ_l is called the *phase shift* in the lth partial wave.

Eq. (5) becomes,

$$\psi(r, \theta) = \sum_l \frac{A_l}{kr}\sin\left(kr - \frac{l\pi}{2} + \delta_l\right)P_l \cos\theta$$

or

$$\Psi(r, \theta) = \sum_l \frac{A_l}{kr}\left\{\frac{e^{i}\left(kr - \frac{l\pi}{2} + \delta_l\right)_{-e} - i\left(kr - \frac{l\pi}{2} + \delta_l\right)}{2i}\right\}\times P_l(\cos\theta) \qquad \text{...(13)}$$

This equation is identical with asymptotic form of Eq. (1), *i.e.*,

$$\Psi(r) = e^{ikz} + f(\theta)\frac{e^{ikr}}{r}$$

$$e^{ikz} = e^{ikr}\cos\theta = \sum_{l=0}^{\infty}(2l+1)i^l j_l(kr)P_l(\cos\theta)$$

$$\xrightarrow[r\to\infty]{} \sum_{l=0}^{\infty}(2l+1)i^l\left(\frac{1}{kr}\right)\sin\left(kr - \frac{1}{2}l\pi\right)P_l(\cos\theta) \qquad \text{...(14)}$$

($\boldsymbol{j_l}$ are spherical Bessel functions)

$$\therefore\ \Psi(r) = \sum_l(2l+1)i^l\cdot\frac{1}{kr}\sin\left(kr - \frac{l\pi}{2}\right)P_l(\cos\theta) + f(\theta)\frac{e^{ikr}}{r}$$

$$\boldsymbol{\Psi(r)} = \sum_l(2l+1)i^l\cdot\frac{1}{kr}\left(\frac{e^{i\left(kr - \frac{l\pi}{2}\right)}_{-e} - i\left(kr - \frac{l\pi}{2}\right)}{2i}\right)P_l(\cos\theta) + r^{-1}f(\theta)e^{ikr} \qquad \text{...(15)}$$

Equations (13) and (15) represent the same function and hence the coefficients of e^{ikr} and e^{-ikr} should be identical.

Equating the coefficients gives us

$$\sum_l A_l\frac{e^{i\left(-\frac{l\pi}{2}+\delta_l\right)}}{2\,i\,kr}P_l(\cos\theta) = \sum_l\frac{(2l+1)\,i^l\,e^{-il\pi/2}}{2\,i\,kr}.P_l(\cos\theta) + r^{-1}.f(\theta) \qquad \text{...(16)}$$

$$\text{and } \sum_l A_l\frac{e^{i\left(-\frac{l\pi}{2}+\delta_l\right)}}{(2ikr)}\times P_l(\cos\theta) = \sum_l\frac{(2l+1)}{2ikr}i^l\,e^{i\,l\pi/2}.P_l(\cos\theta) \qquad \text{...(17)}$$

From Eq. (17), we get

$$A_l = (2l+1)\,i^l\,e^{i\delta_l} \qquad \text{...(18)}$$

Substituting this value of A_l in Eq. (16), we get

$$\sum_l\frac{(2l+1)i^l\,e^{i\delta l}\,e^{i\left(-\frac{l\pi}{2}+\delta_l\right)}}{(2ikr)}P_l(\cos\theta) = \sum_l\frac{(2l+1)\,i^l\,e^{-i\,l\,\pi/2}}{2i\,kr}.P_l(\cos\theta) + \frac{1}{r}f(\theta)$$

$$\text{Put } i^l = e^{il\pi/2}.$$

$$f(\theta) = (2ik)^{-1} \sum_{l=0}^{\infty} (2l+1)\,(e^{2i\delta_l} - 1)\, P_l\,(\cos\theta)$$

$$\text{or} \quad \text{scattering amplitude, } f(\theta) = \frac{1}{k} \sum_{l=0}^{\infty} (2l+1)\, e^{i\,\delta_l} \sin\delta_l\; P_l\,(\cos\theta) \qquad \text{...(19)}$$

This gives $f(\theta)$ as a sum of contributions from partial waves with $l = 0$ to ∞.

The differential scattering cross-section,

$$\frac{d\sigma(\theta)}{d\Omega} = |\,f(\theta)|^2 = \frac{1}{k^2} \left| \sum_{l=0}^{\infty} (2l+1)\, P_l\,(\cos\theta)\, e^{i\delta_l} \sin\delta_l \right|^2 \qquad \text{...(20)}$$

The total scattering cross-section,

$$\sigma = 2\pi \int \frac{\sigma}{} \,.\, \sin\theta\, d\theta = \frac{\pi}{} \sum^{\infty} (2l+1) \sin\ \ \delta \qquad \text{...(21)}$$

A knowledge of the phase shifts δ_l enables us to calculate the total scattering cross-section.

16.10 Transformation from Centre of Mass to Laboratory Frame of Reference

The scattering of particles can be visualised in two kinds of coordinates :

1. *Laboratory frame or system* **(L-system) :** It is that coordinate system in which the bombarded particle (or target) is initially at rest [Fig. 16. 7 (a)].

2. *Centre of mass coordinate system* **(C-system) :** It is that coordinate system in which the centre of mass of two colliding particles is at rest (initially and always) [Fig. 16.7 (b)].

In the centre of mass system, the two particles approach each other with the same momentum.

Relation between Angles in L-system and C-system

(*a*) Consider a particle of mass m_1 moving with velocity v_1 towards the target particle of mass m_2 which is initially at rest (Fig. 16.7 (a)).

The centre of mass moves with volocity

$$v_c = \frac{m_1 v_1}{m_1 + m_2} \qquad \text{...(1)}$$

v' is the speed of the particle of mass m_1 after scattering.

After scattering, the incident particle moves in a direction given by the angle (θ_0, ϕ_0).

The target particle also suffers a recoil and moves in a different direction.

(*b*) In the centre of mass system, the centre of mass is at rest and the two particles will approach each other with velocities $m_2 v_1/(m_1 + m_2)$ and $- m_1 v_1/(m_1 + m_2)$.

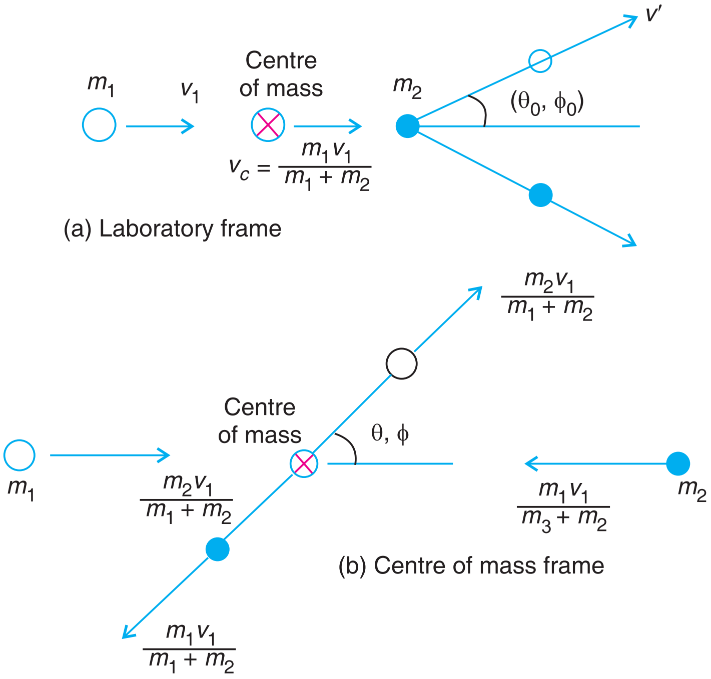

v'
Centre of mass
m_1
v_1
m_2
(θ_0, ϕ_0)
$v_c = \frac{m_1 v_1}{m_1 + m_2}$
(a) Laboratory frame
$\frac{m_2 v_1}{m_1 + m_2}$
Centre of mass
θ, ϕ
m_1
$\frac{m_2 v_1}{m_1 + m_2}$
$\frac{m_1 v_1}{m_3 + m_2}$
m_2
(b) Centre of mass frame
$\frac{m_1 v_1}{m_1 + m_2}$

Fig. 16.7

The total momenta of the two particles before scattering is zero.

Hence they must fly off in opposite directions after scattering (Fig. 16.7 (b)).

In the centre of mass system, the scattering angle is denoted by (θ, ϕ).

As collision is symmetric in ϕ-coordinate,

$$\phi = \phi_0 \qquad \text{...(2)}$$

For elastic scattering, the particles should move away with their original speeds. In order to get back to the laboratory system we must superpose the velocity of the centre of mass.

$$v' \sin\theta_0 = \frac{m_2 v_1}{m_2 + m_1} \sin\theta \qquad \text{...(3)}$$

$$v' \cos\theta_0 = \frac{m_2 v_1}{m_2 + m_1} \cos\theta + v_c \qquad \text{...(4)}$$

$$\therefore \quad \tan\theta_0 = \frac{\sin\theta}{\frac{m_1}{m_2} + \cos\theta} \qquad \text{...(5)}$$

When $m_2 >> m_1$, the centre of mass is almost at rest, $\theta_0 \simeq \theta$ and the two system of coordinates coincide.

Relation between cross-sections in L-system and C-system

Since the laboratory and centre of mass scattering angles are not equal, the angular distribution of the cross-section will be different in the two systems. The relation between the two scattering cross-sections can be obtained by noting that the number of particles scattered in a given element of solid angle must be the same in two systems. Thus

$$\left(\frac{d\sigma}{d\Omega}\right)_0 (\theta_0, \varphi_0) \sin\theta_0 \, d\theta_0 d\phi_0 = \frac{d\sigma}{d\Omega}(\theta, \varphi) \sin\theta \, d\theta \, d\varphi. \qquad \text{...(6)}$$

The azimuthal angles in the two frames are equal : $\varphi_0 = \varphi$.

or $$d\phi_0 = d\phi \qquad \text{...(7)}$$

$$\therefore \quad \left(\frac{d\sigma}{d\Omega}\right)_0 (\theta_0, \varphi_0) = \frac{d\sigma}{d\Omega}(\theta, \varphi).\frac{\sin\theta}{\sin\theta_0}\left|\frac{d\theta}{d\theta_0}\right| \qquad \text{...(8)}$$

From Eq. (5), $$\frac{d\theta}{d\theta_0} = \frac{(m_1/m_2)^2 + 2(m_1/m_2)\cos\theta + 1}{(m_1/m_2)\cos\theta + 1} \qquad \text{...(9)}$$

$$\left(\frac{d\sigma}{d\Omega}\right)_0 (\theta_0, \phi_0) = \frac{[(m_1/m_2)^2 + 2\,(m_1/m_2)\cos\theta + 1]^{3/2}}{|1 + (m_1\,/\,m_2)\cos\theta\,|}\;\frac{d\sigma}{d\Omega}(\theta, \varphi) \qquad \text{...(10)}$$

It is $\left(\frac{d\sigma}{d\Omega}\right)(\theta, \phi)$ which is obtained from the theory.

Eq. (10) tells us how to transform it to the laboratory system which should be compared with the experimental data.

16.11 Reduction of the Two-Body Problem into One-Body Problem

(*i*) We first give the simplification of the problem according to *classical mechanics*. If a system of particles is not subject to any external forces (i.e., forces other than the mutual interactions of the particles themselves) then the centre of mass of the system moves with constant velocity $\boldsymbol{v}_c$. In a reference frame which moves along with the same velocity $\boldsymbol{v}_c$, the centre of mass of the particle-system appears to remain at rest. Any such reference frame is called a *centre of mass frame.* In the particular case of a two-body system, the motion as seen from the centre of mass frame is especially simple since, with the centre of mass remaining fixed, the instantaneous position of either particle completely determines that of the other. The two-body problem then becomes effectively a one-body problem.

(*ii*) Simplification of the Problem in Quantum Mechanics

For a two body problem, whenever the potential energy depends only on the magnitude of the distance between the two particles, the problem can always be reduced to a one body problem along with a uniform translational motion of the centre of mass.

Consider a two-particle system with the Hamiltonian

$$H = -\frac{\hbar^2}{2m_1}\nabla_1^2 - \frac{\hbar^2}{2m_2}\nabla_2^2 + V\,(\boldsymbol{x}_1 - \boldsymbol{x}_2) \qquad \text{...(1)}$$

The subscripts 1 and 2 refer to the two particles.

We define the *centre of mass* and *relative* co-ordinate variables as

$$\boldsymbol{X} = -\frac{m_1\boldsymbol{x}_1 + m_2\boldsymbol{x}_2}{m_1 + m_2} \quad \text{and} \quad \boldsymbol{x} = \boldsymbol{x}_1 - \boldsymbol{x}_2 \qquad \text{...(2)}$$

We denote the corresponding gradient operators by

$$\nabla_X \equiv \left(\frac{\partial}{\partial X}, \frac{\partial}{dY}, \frac{\partial}{\partial Z}\right) \text{ and } \nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{dy}, \frac{\partial}{\partial z}\right) \qquad \text{...(3)}$$

Both $\nabla_1 \equiv (\partial/\partial x_1, \partial/\partial y_1, \partial/\partial z_1)$ and ∇_2 can be expressed in terms of these.

$$\frac{\partial}{\partial x_1} = \frac{\partial X}{\partial x_1}\frac{\partial}{\partial X} + \frac{\partial x}{\partial x_1}\frac{\partial}{\partial x} = \frac{m_1}{m_1 + m_2}\frac{\partial}{\partial X} + \frac{\partial}{\partial x}$$

$$\frac{\partial^2}{\partial x_1^2} = \left(\frac{m_1}{m_1 + m_2}\right)^2 \frac{\partial^2}{\partial X^2} + \frac{2m_1}{m_1 + m_2}\frac{\partial^2}{\partial X \partial x} + \frac{\partial^2}{\partial x^2}, \text{ etc.} \qquad \text{...(4)}$$

$$\nabla_1^2 = \left(\frac{m_1}{m_1+m_2}\right)^2 \nabla_X{}^2 + \frac{2m_1}{m_1+m_2}\nabla_X . \nabla + \nabla^2 \qquad \text{...(5)}$$

$$\nabla_2^2 = \left(\frac{m_2}{m_1+m_2}\right)^2 \nabla_X{}^2 - \frac{2m_2}{m_1+m_2}\nabla_X . \nabla + \nabla^2 \qquad \text{...(6)}$$

On substituting Eqs. (5) and (6) into Eq. (1), it becomes

$$H = -\frac{\hbar^2}{2M}\nabla_X{}^2 - \frac{\hbar^2}{2\mu}\nabla^2 + V(\boldsymbol{x}) \qquad \text{...(7)}$$

Here, $$M = m_1 + m_2,\ \mu = \frac{m_1 m_2}{m_1 + m_2} \qquad \text{...(8)}$$

$\mu \equiv m_1 m_2/(m_1 + m_2)$ is called the *reduced mass* of the system.

In this form, H is sum of two commuting parts depending on independent sets of variables.

(*i*) The first part, $H_{cm} = -(\hbar^2/2M)\,\nabla_X^2$ describes the free motion of the centre of mass. It represents the kinetic energy of the system (with total mass M) as a whole.

(*ii*) The second part, $H_{rel} = -\dfrac{\hbar^2}{2\mu}\nabla^2 + V(\boldsymbol{x})$ depends on the relative variable $\boldsymbol{x}$ only. It is identical in form with the Hamiltonian of a single particle of mass μ moving in a potential $V(\boldsymbol{x})$.

A complete set of eigen functions of H is now given by

$$(2\pi)^{-3/2} \exp\,(\boldsymbol{iK.X})\,\Psi_n\,(\boldsymbol{x}) \qquad \text{...(9)}$$

(for all $\boldsymbol{K}$, n)

(*i*) The factor $(2\pi)^{-3/2} \exp\,[iK.X]$ describes the motion of the centre of mass with momentum $\hbar\boldsymbol{K}$.

(*ii*) $\Psi_n(\boldsymbol{x})$ are eigen functions of H_{rel} given by

$$-\frac{\hbar^2}{2M}\nabla_X{}^2\,[(2\pi)^{-3/2} \exp\,(i\boldsymbol{K}.\boldsymbol{X})] = \frac{\hbar^2 K^2}{2M}[(2\pi)^{-3/2} \exp\,(i\boldsymbol{K}.\boldsymbol{X})] \qquad \text{...(10)}$$

$$\left[-\frac{\hbar^2}{2\mu}\nabla^2 + V(\boldsymbol{x})\right]\Psi_n(\boldsymbol{x}) = E_n \Psi n(\boldsymbol{x}) \qquad \text{...(11)}$$

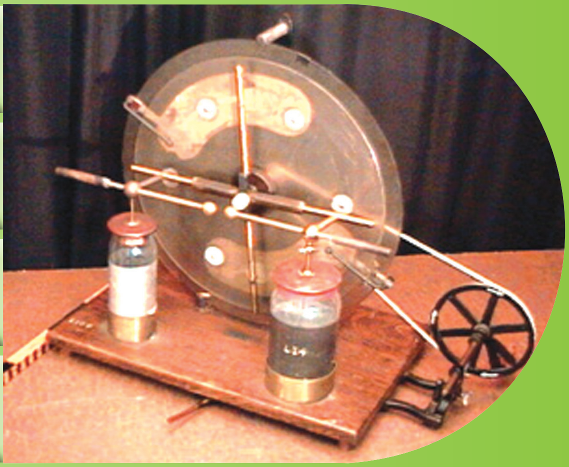
Thus the problem reduce essentially to that of solving the single-particle Schrodinger equation.

EXERCISE

1. What do you mean by (*i*) differential scattering cross-section (*ii*) total scattering cross-section and (*iii*) scattering amplitude ? (*Rohilkhand, 1997*)
2. Show that the differential scattering cross-section is equal to the square of the scattering amplitude.
3. Explain the method of partial waves to calculate the phase shifts and scattering amplitude. (*Agra, 1982*)
4. Illustrate with diagrams elastic scattering in laboratory and centre of mass systems. (*Agra, 1983*)

CHAPTER

17



THREE DIMENSIONAL HARMONIC OSCILLATOR

17.1 Three Dimensional Harmonic Oscillator—(Spherically Symmetric Case)

A three dimensional harmonic oscillator consists of a particle bound to origin by a force proportional to the displacement r from the fixed point, *i.e.*,

$$F = -kr.$$

Here, k is force constant.

We consider the special case in which frequency of oscillator along three axes X, Y and Z are same, *i.e.*,

$$\nu_x = \nu_y = \nu_z = \nu \text{ (say)}$$

$$\therefore \quad k = m\omega^2 = m(2\pi\nu)^2 = 4\pi^2\nu^2 m. \qquad \text{...(1)}$$

Potential energy of oscillator assuming zero potential energy at $r = 0$ is

$$V = -\int F\,dr = \int kr\,dr = \frac{1}{2}kr^2$$

But

$$r^2 = x^2 + y^2 + z^2$$

$$\therefore \quad V(r) = \frac{1}{2}kr^2 = \frac{1}{2}k(x^2 + y^2 + z^2).$$

The Schroedinger equation for this system is

$$\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2} + \frac{2m}{\hbar^2}\left[E - \frac{1}{2}k(x^2 + y^2 + z^2)\right]\psi = 0$$

or

$$\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2} + 2\sqrt{\frac{m}{\hbar^2}}\sqrt{\frac{mk}{\hbar^2}}$$

$$\times\left[\frac{E}{\sqrt{k}} - \frac{1}{2}\sqrt{k}(x^2 + y^2 + z^2)\right]\psi = 0.$$

or

$$\frac{1}{\sqrt{\frac{mk}{\hbar^2}}}\left\{\frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2}\right\}$$

$$+\left\{2\sqrt{\frac{m}{\hbar^2 k}}E-\sqrt{\frac{mk}{\hbar^2}}(x^2+y^2+z^2)\right\}\psi = 0 \quad \text{...(2)}$$

Let $$\sqrt{\frac{mk}{\hbar^2}} = \alpha^2 \text{ and } 2E\sqrt{\frac{m}{\hbar^2 k}} = \lambda \quad \text{...(3)}$$

$\therefore$ $$\frac{1}{\alpha^2}\left\{\frac{\partial^2 \psi}{\partial x^2}+\frac{\partial^2 \psi}{\partial y^2}+\frac{\partial^2 \psi}{\partial z^2}\right\}+[\lambda-\alpha^2(x^2+y^2+z^2)]\psi = 0 \quad \text{...(4)}$$

Let $$q_x = \alpha x,\ q_y = \alpha y \text{ and } q_z = \alpha z \quad \text{...(5)}$$

$\therefore$ $$\frac{\partial^2 \psi}{\partial q_x^2}+\frac{\partial^2 \psi}{\partial q_y^2}+\frac{\partial^2 \psi}{\partial q_z^2}+[\lambda-(q_x^2+q_y^2+q_z^2)]\psi = 0 \quad \text{...(6)}$$

This equation can be solved by the method of separation of variables.
Let

$$\psi = Q_x(q_x)\, Q_y(q_y)\, Q_z(q_z), \quad \text{...(7)}$$

Here, Q_x is function of q_x, Q_y is function of q_y and Q_z is function of q_z only.
On substituting Eq. (7) in Eq. (6) and simplifying, we get

$$\frac{\partial^2 Q_x}{\partial q_x^2}+(\lambda_x-q_x^2)Q_x = 0 \quad \text{...(8)}$$

$$\frac{\partial^2 Q_y}{\partial q_y^2}+(\lambda_y-q_y^2)Q_y = 0 \quad \text{...(9)}$$

$$\frac{\partial^2 Q_z}{\partial q_z^2}+(\lambda_z-q_z^2)Q_z = 0 \quad \text{...(10)}$$

Here, the constants λ_x, λ_y and λ_z are related by

$$\lambda_x+\lambda_y+\lambda_z = \lambda \quad \text{...(11)}$$

Each of Eqs. (8), (9) and (10) is the differential equation for a one-dimensional harmonic oscillator (Section 11.13). By analogy, the normalised wave functions are given by

$$Q_x(q_x) = \frac{1}{\left[2^{n_x} n_x!\sqrt{\pi}\right]^{1/2}} e^{-q_x^2/2} H_{n_x}(q_x) \quad \text{...(12)}$$

$$Q_y(q_y) = \frac{1}{\left[2^{n_y} n_y!\sqrt{\pi}\right]^{1/2}} e^{-q_y^2/2} H_{n_y}(q_y) \quad \text{...(13)}$$

$$Q_z(q_z) = \frac{1}{\left[2^{n_z} n_z!\sqrt{\pi}\right]^{1/2}} e^{-q_z^2/2} H_{n_z}(q_z) \quad \text{...(14)}$$

λ_x, λ_y and λ_z are restricted to the values

$$\lambda_x = 2n_x+1,\ \lambda_y = 2n_y+1 \text{ and } \lambda_z = 2n_z+1 \quad \text{...(15)}$$

Here, n_x, n_y and n_z are *non-negative integers.*

The *complete normalised wave function* is given by

$$\psi = Q_x Q_y Q_z = \frac{1}{\left[2^{(n_x+n_y+n_z)} n_x!n_y!n_z!\pi^{3/2}\right]^{1/2}} e^{-1/2(q_x^2+q_y^2+q_z^2)}$$

$$\times H_{n_x}(q_x)H_{n_y}(q_y)H_{n_z}(q_z). \quad \text{...(16)}$$

The *eigen values* are given by

$$E_n = \left(n_x + n_y + n_z + \frac{3}{2}\right)\hbar\omega = \left(n + \frac{3}{2}\right)\hbar\omega \qquad \text{...(17)}$$

Here, $n = n_x + n_y + n_z$ is called *total quantum number.*

Energy of this system depends only on the total quantum number, n.

The states of energy E_n are, therefore, *degenerate*. The degree of the degeneracy is $\frac{1}{2}[(n+1)(n+2)]$. Fig. 17.1 shows the energy levels and degree of degeneracy for three dimensional isotropic harmonic oscillator.

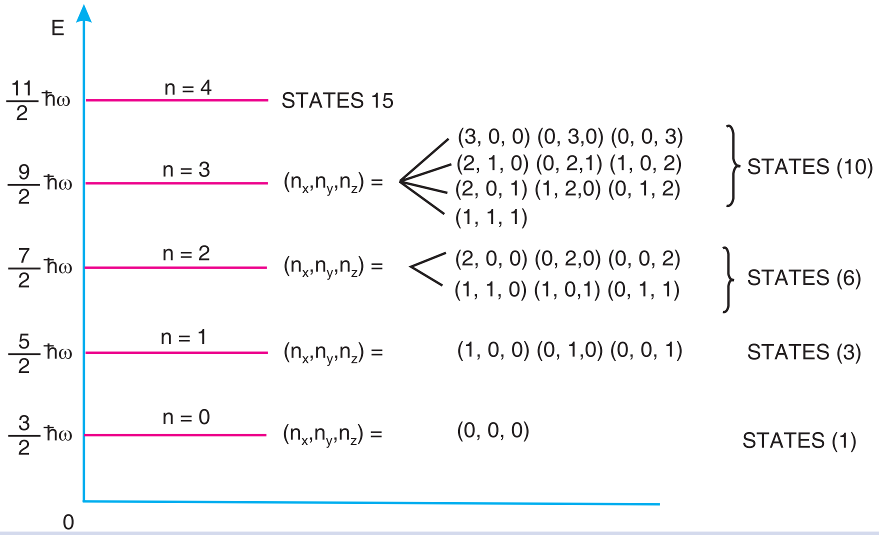

E
$\frac{11}{2}\hbar\omega$
n = 4
STATES 15
$\frac{9}{2}\hbar\omega$
n = 3
(n_x,n_y,n_z) =
(3, 0, 0) (0, 3,0) (0, 0, 3)
(2, 1, 0) (0, 2,1) (1, 0, 2)
(2, 0, 1) (1, 2,0) (0, 1, 2)
(1, 1, 1)
STATES (10)
$\frac{7}{2}\hbar\omega$
n = 2
(n_x,n_y,n_z) =
(2, 0, 0) (0, 2,0) (0, 0, 2)
(1, 1, 0) (1, 0,1) (0, 1, 1)
STATES (6)
$\frac{5}{2}\hbar\omega$
n = 1
(n_x,n_y,n_z) =
(1, 0, 0) (0, 1,0) (0, 0, 1)
STATES (3)
$\frac{3}{2}\hbar\omega$
n = 0
(n_x,n_y,n_z) =
(0, 0, 0)
STATES (1)
0

Fig. 17.1

EXERCISE

1. Find the solution of three dimensional harmonic oscillator in rectangular coordinates. (*Garhwal*, 1992)
2. Solve the problem of a three dimensional isotropic harmonic oscillator in quantum mechanics. (*Rohilkhand*, 1979)

CHAPTER

18

IDENTICAL PARTICLES AND SPIN

AT A GLANCE

18.1 Identical Particles
18.2 Symmetric and Antisymmetric Wave Functions
18.3 Construction of Symmetric and Antisymmetric Wave Functions from Unsymmetrised Functions
18.4 Distinguishability of Identical Particles
18.5 The Pauli Exclusion Principle
18.6 Spin
18.7 Pauli's Spin Matrices for Electron
18.8 Commutation Relations Satisfied by Three Components σ_x, σ_y and σ_z
18.9 Two Component Wave Functions

18.1 Identical Particles

Identical particles in a system are regarded as those particles which when interchanged in the system will not make any change in it. These particles can be distinguished from one another, only when their wave packets do not overlap. Similarly if the particles have spins which are aligned in different positions in the process of any interaction, they can be identified from one another as a result of such interactions.

Thus the word "identical" in quantum mechanics is to describe particles that can be substituted for each other under the most general possible circumstances with no change in physical situation.

For example, the component of spin along some particular axis is assumed not to change during elastic collisions. It means the particles can be distinguished if they have different spin components.

18.2 Symmetric and Antisymmetric Wave Functions

The Schroedinger equation for *n* identical particles is

$$H(1,2,3....n)\,\psi(1,2,3....n;t) = i\hbar\frac{\partial}{\partial t}\psi(1,2,3....n,t) \qquad ..(1)$$

Here each of the numbers represents all the coordinates (positional and spin) of one of the particles. The Hamiltonian H is symmetrical in its arguments due to the identity of particles. The identity of particles means that they can be substituted for each other without changing H or indeed any other observable.

Now we define symmetric and antisymmetric wave functions.

(*i*) Symmetric wave function ψ_s. *A wave function is symmetric if the interchange of any pair of particles among its arguments leaves the wave function unchanged.*

(*ii*) Antisymmetric wave function ψ_A. *A wave function is antisymmetric if the interchange of any pair of particles among its arguments changes the sign of the wave function.*

If P is an exchange operator, then we must have

$$P\psi_s(1, 2) = \psi_s(2, 1),$$
$$P\psi_A(1, 2) = -\psi_A(2, 1).$$

18.3 Construction of Symmetric and Antisymmetric Wave Functions from Unsymmetrised Functions

We now show how ψ_s or ψ_A can be constructed from a general unsymmetrized solution ψ of Eq. (1) [Sec. 18.2]. If the arguments of the wave function ψ are permuted in any way, the resulting wave function is also a solution of Eq. (1). This is because the same permutation applied throughout Eq. (1) does not impair its validity, since it corresponds simply to a relabeling of the particles. Since H is symmetric, the permuted H is the same as the original H, and the resulting equation is the same as Eq. (1) for the permuted ψ. In this way $n!$ solutions can be obtained from any one solution, each of which corresponds to one of the $n!$ permutations of the n arguments of ψ. Evidently any linear combination of these functions is also a solution of the wave equation (1).

The sum of all these functions is a symmetric (unnormalised) wave function ψ_s, since the interchange of any pair of particles changes any one of the component functions into another of them and the latter into the former, leaving the entire wave function unchanged.

An antisymmetric unnormalised wave function can be constructed by adding together all the permuted wave functions that arise from the original solution by means of an even number of interchanges of pairs of particles, and subtracting the sum of all the permuted wave functions that airse by means of an odd number of interchanges of pairs of particles in the original solution.

For a system of 3 particles, the Schroedinger's equation is

$$H(1, 2, 3)\,\psi(1, 2, 3) = E\psi(1, 2, 3) \qquad ...(1)$$

This equation has $3! = 6$ solutions corresponding to the same eigen values E.

The six possible functions obtained by exchanging the indices of the particles are

$$\psi(1, 2, 3), \psi(2, 3, 1), \psi(3, 2, 1), \psi(1, 3, 2), \psi(2, 1, 3), \psi(3, 1, 2).$$

Out of these six functions, those arising by an even number of interchanges of the pairs of particles in original wave function $\psi(1, 2, 3)$ are

$$\psi(1, 2, 3), \psi(2, 3, 1), \psi(3, 1, 2),$$

The functions arising by odd number of interchanges of pairs of particles in original function $\psi(1, 2, 3)$ are

$$\psi(2, 1, 3), \psi(1, 3, 2), \psi(3, 2, 1).$$

So the *symmetric wave function* can be written as

$$\psi_S = \psi(1,2,3)+\psi(2,3,1)+\psi(3,2,1)+\psi(1,3,2) + \psi(2,1,3)+\psi(3,1,2), \quad \text{...(2)}$$

The *antisymmetric wave function* is

$$\psi_A = \psi(1,2,3)+\psi(2,3,1)+\psi(3,1,2)-\{\psi(1,3,2) + \psi(2,1,3)+\psi(3,2,1)\}. \quad \text{...(3)}$$

Here ψ_S and ψ_A are unnormalised wave functions.

18.4 Distinguishability of Identical Particles

Two identical particles will be distinguishable from each other if the sum of probabilities of the individual wave functions in two states is equal to the probability derived by the symmetrised wave function as given by the following equations:

$$\begin{aligned}\psi_S &= \psi(1,2)+\psi(2,1)\\ \psi_A &= \psi(1,2)-\psi(2,1)\\ |\psi(1,2)|^2+|\psi(2,1)|^2 &= |\{\psi(1,2)\pm\psi(2,1)\}|^2\\ &= |\psi(1,2)|^2+|\psi(2,1)|^2\\ &\quad \pm 2\,Re\,[\psi(1,2)\,\psi^*(2,1)] \quad \text{...(1)}\end{aligned}$$

Here, Re denotes the real part of $\{\psi(1,2)\,\psi^*(2,1)\}$.

This is possible only when overlap of wave-functions $\psi(1,2)$ and $\psi(2,1)$ is zero

or $2\,Re\,\{\psi(1,2)\,\psi^*(2,1)\} = 0.$

Thus when the co-ordinates (space and spin) of two particles are not the same between exchange degenerate functions, the interference term, *i.e.*, $2\,Re\,\psi(1,2)\,\psi^*(2,1)$ becomes zero and particle co-ordinates do not overlap.

18.5 The Pauli Exclusion Principle

Schroedinger's time dependent equation for a system of N-particles is

$$i\hbar\frac{\partial\psi}{\partial t}(1,2,3,r\ldots s\ldots N;t) = \hat{H}(1,2,\ldots N)\,\psi(1,2,3\ldots r\ldots s\ldots N;t) \quad \text{...(1)}$$

As Hamiltonian does not involve time, the stationary solution can be found.

Eq. (1) can be written as

$$[\hat{H}(1,2,3\ldots N)-E]\,\psi(1,2,3\ldots r\ldots s\ldots N)=0 \quad \text{...(2)}$$

A useful zero-order approximation can be obtained by neglecting the interactions between the particles that make up the system under consideration. The approximate (unperturbed) Hamiltonian is the sum of equal Hamiltonian functions for the separate particles.

$$H_0(1,2,3\ldots N) = H_0'(1)+H_0'(2)+\ldots+H_0'(N), \quad \text{...(3)}$$

Here, $H_0(1,2,3\ldots N)$ is unperturbed Hamiltonian for the complete system. $H'_0(1)$ etc., are the unperturbed Hamiltonian for the individual particles.

The approximate energy eigenfunction is a product of one-particle eigenfunctions of H'_0.

$$\psi(1,2,3)\ldots N) = \psi_{n1}(1)\,\psi_{n2}(2)\ldots\psi_{nN}(N) \quad \text{...(4)}$$

Here, n_1, n_2, n_N indicate the totality of the quantum numbers required to represent the state. If E represents the total energy of the system,

$$E = E_{n1}+E_{n2}+\ldots.\,E_{nN}, \quad \text{...(5)}$$

Moreover, the wave functions are such that they obey

$$H_0'(1)\,\psi_{n1}(1) = E_{n1}\,\psi_{n1}(1) \text{ etc.} \qquad ...(6)$$

Thus for Fermi particles, the antisymmetric state which only describes the actual state of affairs can be written as follows:

$$\psi_A(1, 2, 3 \,.....\, N) = \frac{1}{\sqrt{N!}}\begin{vmatrix} \psi_{n1}(1)\ \psi_{n1}(2) \,.... & \psi_{n1}(N) \\ \psi_{n2}(1)\ \psi_{n2}(2) \,.... & \psi_{n2}(N) \\ ...\ ...\ ...\ ... & \\ ...\ ...\ ...\ ... & \\ \psi_{nN}(1)\ \ \psi_{nN}(2)...... & \psi_{nN}(N) \end{vmatrix} \qquad ...(7)$$

Here ψ_A is solution of the equation

$$H_0(1, 2, 3 \,...N)\,\psi_A(1, 2, \,...\, N) = E\,\psi_A(1, 2 \,...\, N) \qquad ...(8)$$

Eq. (7) suggests that if two ψ_s'are the same, the state does not exist. Thus the unperturbed Hamiltonian H_0 has no solutions for which there is more than one electron in any one of the states 1,2, 3 ..., N. This result is known as the *Pauli's exclusion principle.* It states that *no two particles described by antisymmetric wave functions* (or obeying Fermi-Dirac Statistics) *can exist in same quantum state.*

18.6 Spin

Spin is an intrinsic property of particles. This property was deduced from the Stern-Gerlach experiment. From other observations we conclude that spin is a property of elementary particles, such as protons, neutrons, and electrons. Spin is a vector quantity. The characteristic values of the component of the spin in any prescribed direction are given by $\hbar/2$ and $-\hbar/2$. Moreover, the spin is measured in unit of $\hbar$ and the characteristic values are $\frac{1}{2}$ and $-\frac{1}{2}$. This type of behavior can be described by a spin quantum number $s = \frac{1}{2}$.

18.7 Pauli's Spin Matrices for Electron

Like orbital angular momentum operators L_x, L_y, L_z the spin operators S_x, S_y and S_z to be associated with the components of spin angular momentum satisfy the commutation relations,

$$\left.\begin{aligned} [S_x, S_y] &= S_x S_y - S_y S_x = i\hbar S_z \\ [S_y, S_z] &= S_y S_z - S_z S_y = i\hbar S_x \\ [S_z, S_x] &= S_z S_x - S_x S_z = i\hbar S_y \end{aligned}\right] \qquad ...(1)$$

We consider the electron, with spin $= \frac{1}{2}$. Then, according to Uhlenbeck and Goudsmit hypothesis, each of operators, S_x, S_y and S_z must have just two eigen values $\frac{1}{2}\hbar$ and $-\frac{1}{2}\hbar$. Now introduce new auxiliary operators σ_x, σ_y and σ_z such that

$$\left.\begin{aligned} S_x &= \frac{1}{2}\hbar\sigma_x \\ S_y &= \frac{1}{2}\hbar\sigma_y \\ S_z &= \frac{1}{2}\hbar\sigma_z \end{aligned}\right] \qquad ...(2)$$

The following properties of σ's may be noted:

Since the eigen values of each S are to be just $\frac{1}{2}\hbar$ and $-\frac{1}{2}\hbar$, the eigen values of each σ must be 1 and – 1. Each of the operators σ_x^2, σ_y^2 and σ_z^2 must therefore have only the eigen value 1 and such operator is only unit operator. Therefore,

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = 1 \qquad ...(3)$$

According to Eqs. (2) and (3), the commutation rules satisfied by σ's must be

$$\left.\begin{aligned}[\sigma_x, \sigma_y] &= \sigma_x\,\sigma_y - \sigma_y\,\sigma_x = 2i\sigma_z\\ [\sigma_y, \sigma_z] &= \sigma_y\,\sigma_z - \sigma_z\,\sigma_y = 2i\sigma_x\\ [\sigma_z, \sigma_x] &= \sigma_z\,\sigma_x - \sigma_x\,\sigma_z = 2i\sigma_y\end{aligned}\right\} \qquad ...(4)$$

Now

$$\begin{aligned}2i\,(\sigma_x\sigma_y + \sigma_y\sigma_x) &= (2i\sigma_x)\,\sigma_y + \sigma_y\,(2i\sigma_x)\\ &= (\sigma_y\sigma_z - \sigma_z\sigma_y)\,\sigma_y + \sigma_y\,(\sigma_y\sigma_z - \sigma_z\sigma_y)\\ &= -\sigma_z\sigma_y{}^2 + \sigma_y{}^2\,\sigma_z\\ &= -\sigma_z + \sigma_z\\ &= 0.\end{aligned}$$

Hence $\sigma_x\sigma_y = -\sigma_y\sigma_x$.

∴ σ_x and σ_y anticommute.

Similarly, any two of the σ's anticommute as

$$\left.\begin{aligned}\sigma_x\sigma_y + \sigma_y\sigma_x &= 0\\ \sigma_y\sigma_z + \sigma_z\sigma_y &= 0\\ \sigma_z\sigma_x + \sigma_x\,\sigma_z &= 0\end{aligned}\right\} \qquad ...(5)$$

Finally from Eqs. (4) and (5) we have

$$\left.\begin{aligned}\sigma_x\sigma_y &= i\sigma_z, \qquad ...(6a)\\ \sigma_y\sigma_z &= i\sigma_x, \qquad ...(6b)\\ \sigma_z\sigma_x &= i\sigma_y, \qquad ...(6c)\end{aligned}\right\}$$

Each σ has two eigen values. So (2 × 2) matrix may be expected to fulfil the purpose. We begin by associating with σ_z, the simplest (2 × 2) matrix having the eigen values 1 and – 1.

$$\sigma_z = \begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix} \qquad ...(7)$$

$$\sigma_x\,\sigma_z = \begin{bmatrix}a & b\\ c & d\end{bmatrix}\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix} = \begin{bmatrix}a & -b\\ c & -d\end{bmatrix} \qquad ...(8)$$

and

$$\sigma_z\,\sigma_x = \begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}\begin{bmatrix}a & b\\ c & d\end{bmatrix} = \begin{bmatrix}a & b\\ -c & -d\end{bmatrix} \qquad ...(9)$$

But σ_x and σ_z anticommute.

∴

$$\sigma_x\,\sigma_z + \sigma_z\,\sigma_x = \begin{bmatrix}a & -b\\ c & -d\end{bmatrix} + \begin{bmatrix}a & b\\ -c & -d\end{bmatrix} = 0$$

i.e.,

$$\begin{bmatrix}2a & 0\\ 0 & -2d\end{bmatrix} = 0.$$

This yields $a = d = 0.$

So every matrix that anticommutes with (7) as σ_x and σ_y must have the form

$$\sigma_x = \begin{bmatrix} 0 & b \\ c & 0 \end{bmatrix}. \qquad \text{...(10)}$$

The eigen values of (10) are $\pm\sqrt{(bc)}$, so that if they are to be 1 and – 1, we must set $bc = 1$. And simplest possibility is to take $b = c = 1$.

$$\therefore \qquad \sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

It then follows from (6*c*) that the matrix to be associated with σ_y is

$$\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$$

Hence complete list of σ's becomes

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \qquad \text{...(11)}$$

These matrices σ_x, σ_y, and σ_z are called *Pauli spin matrices* associated with the components of spin angular momentum

$$S_x = \frac{1}{2}\hbar\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad S_y = \frac{1}{2}\hbar\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad S_z = \frac{1}{2}\hbar\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \qquad \text{...(12)}$$

Thus we have

$$\sigma^2 = \sigma_x^2 + \sigma_y^2 + \sigma_z^2 = 3$$

$$S^2 = S_x^2 + S_y^2 + S_z^2 = \frac{\hbar^2}{4}(\sigma_x^2 + \sigma_y^2 + \sigma_z^2) = \frac{3}{4}\hbar^2$$

18.8 Commutation Relations Satisfied by Three Components σ_x, σ_y and σ_z

(*i*) We have

$$[\sigma_x, \sigma_y] = \sigma_x\sigma_y - \sigma_y\sigma_x.$$

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

$$[\sigma_x, \sigma_y] = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} - \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} - \begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix} = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} + \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}$$

$$= 2\begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} = 2i\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

$$= 2i\sigma_z.$$

(*ii*)

$$[\sigma_y, \sigma_z] = \sigma_y\sigma_z - \sigma_z\sigma_y.$$

$$= \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} - \begin{bmatrix} 0 & -i \\ -i & 0 \end{bmatrix}$$

$$
\begin{aligned}
&= \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} + \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} \\
&= 2\begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix} = 2i\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\
&= 2i\sigma_x.
\end{aligned}
$$

(*iii*)

$$
\begin{aligned}
[\sigma_z, \sigma_x] &= \sigma_z\sigma_x - \sigma_x\sigma_z. \\
&= \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} - \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \\
&= 2\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} = -2\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\
&= 2i^2\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = 2i\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \\
&= 2i\sigma_y.
\end{aligned}
$$

(*iv*) Commutation relation satisfied by σ^2 and its components σ_x, σ_y and σ_z

$$
\begin{aligned}
&\text{We know that} & \sigma^2 &= \sigma_x^2 + \sigma_y^2 + \sigma_z^2. \\
&\text{Let us take} & [\sigma^2, \sigma_x] &= [\sigma_x^2 + \sigma_y^2 + \sigma_z^2; \sigma_x] \\
& & &= [\sigma_x^2, \sigma_x] + [\sigma_y^2, \sigma_x] + [\sigma_z^2, \sigma_x] \\
&\text{But} & [\sigma^2, \sigma_x] &= \sigma_x^3 - \sigma_x^3 = 0. \\
&\text{So} & [\sigma^2, \sigma_x] &= [\sigma_y^2, \sigma_x] + [\sigma_z^2, \sigma_x] \\
& & &= [\sigma_y\,\sigma_y, \sigma_x] + [\sigma_z\sigma_z, \sigma_x]
\end{aligned}
$$

We know that

$$
\begin{aligned}
& & [ab, c] &= a\,[b, c] + [a, c]\,b \\
&\text{So} & [\sigma^2, \sigma_x] &= \sigma_y\,[\sigma_y, \sigma_x] + [\sigma_y, \sigma_x]\,\sigma_y + \sigma_z\,[\sigma_z, \sigma_x] + [\sigma_z, \sigma_x]\,\sigma_z \\
& & &= \sigma_y\,(-2i\sigma_z) + (-2i\sigma_z)\,\sigma_y + \sigma_z\,(2i\sigma_y) + (2i\sigma_y)\,\sigma_z \\
& & &= -2i\,\sigma_y\sigma_z - 2i\sigma_z\sigma_y + 2i\sigma_z\sigma_y + 2i\sigma_y\sigma_z \\
& & &= 0 \\
&\therefore & [\sigma^2, \sigma_x] &= 0 \\
&\text{Similarly,} & [\sigma^2, \sigma_y] &= 0 \\
& & [\sigma^2, \sigma_z] &= 0.
\end{aligned}
$$

Thus σ^2 commutes with each component σ_x, σ_y and σ_z.

18.9 Two Component Wave Functions

Pauli operators σ_x, σ_y, σ_z are 2 × 2 matrices. Therefore, the Pauli operands must be two-component column matrices. Pauli theory retains in essence the Schroedinger operator for dynamical variables having classical analogues. So the components of these column symbols must be function of (x, y, z) or (r, θ, ϕ) and so on. Thus simplest Pauli operand ψ has the form

$$\psi = \begin{pmatrix} \psi_1 (x, x, z) \\ \psi_2 (x, y, z) \end{pmatrix} \qquad ...(1a)$$

or
$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}. \qquad ...(1b)$$

A Pauli wave-function ψ is said to be *well behaved* if at least one of its components is non-zero and if each of its non-zero components is well behaved in the sense of Schrodinger's theory. The complex conjugate of Pauli wave function corresponding to (1) is the row matrix

$$\psi^* = [\psi_1{}^*, \psi_2{}^*] \qquad ...(2)$$

$$\therefore \qquad \psi^*\psi = [\psi_1^*, \psi_2^*]\begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} = \psi_1{}^*\psi_1 + \psi_2{}^*\psi_2 \qquad ...(3)$$

The Pauli wave function is said to be normalised if

$$\int \psi^* \psi \, d\tau = 1 \qquad ...(4a)$$

or
$$\int (\psi_1^* \psi_1 + \psi_2^* \psi_2) \, d\tau = 1 \qquad ...(4b)$$

Here, the integration extends over the entire three-dimensional space.

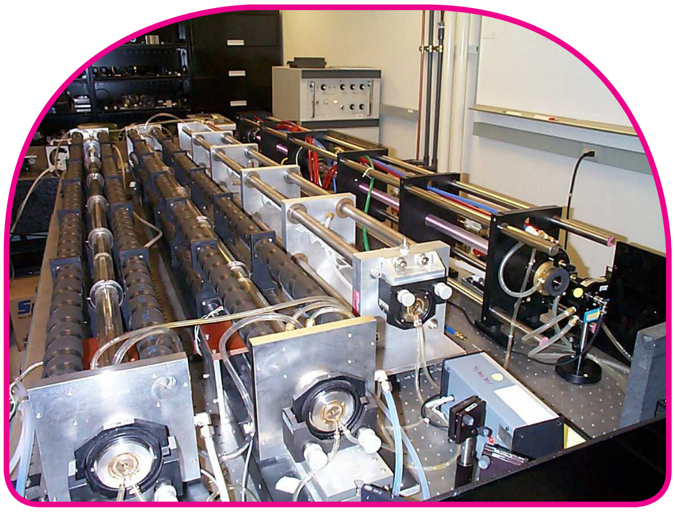
EXERCISE

1. What are symmetric and antisymmetric wave functions? Show how they lead to the Pauli exclusion principle. (*Rohilkhand*, 1984)
2. Construct symmetric and antisymmetric wave functions for three spinless identical particles from unsymmetrised wave function. Show that the antisymmetric wave function obeys Pauli exclusion principle. (*Agra*, 1976)
3. What are Pauli spin operators? Express Pauli spin functions in the form of 2 × 2 matrices.
4. Explain why Pauli introduced a set of 2 × 2 spin matrices and obtain the commutation relations satisfied by the three components of the spin vector. (*Agra*, 1970)

PART – IV

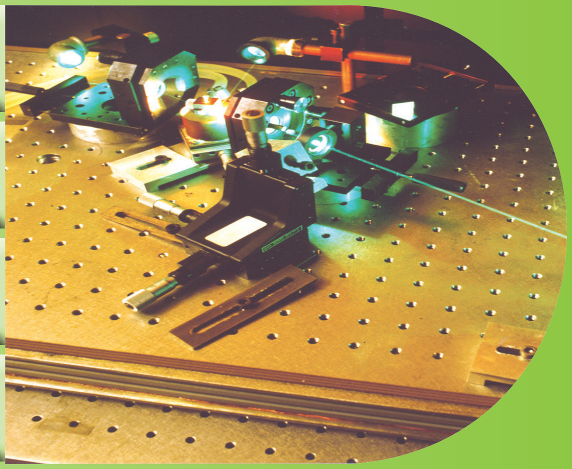
MOLECULAR PHYSICS

19. Lasers, Molecular Spectra and Raman Effect
20. Fine Structure of H_α Line
21. Electronic Spectra of Molecules
22. Atomic Spectra
23. Molecular Spectra
24. Spectroscopic Techniques
25. Spectra of Alkali Elements
26. X-Ray Spectra



CHAPTER

19



LASERS, MOLECULAR SPECTRA AND RAMAN EFFECT

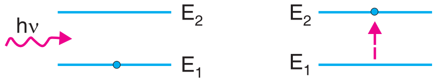
AT A GLANCE

19.1	Induced Absorption, Spontaneous Emission and Stimulated Emission	**19.2**	The Ruby Laser
19.3	Helium-neon Laser	**19.4**	Semiconductor Laser
19.5	Properties of a Laser Beam	**19.6**	Molecular spectra
19.7	Theory of the Origin of Pure Rotational Spectrum of a Molecule	**19.8**	Theory of the Origin of the Vibration-rotation Spectrum of a Molecule
19.9	Electronic Spectra of Molecules	**19.10**	Rayleigh's Scattering
19.11	Discovery	**19.12**	Experimental Study of Raman Effect
19.13	Quantum Theory of Raman Effect	**19.14**	Applications
19.15	Laser Raman Spectroscopy	**19.16**	Rabi's Molecular Beam Magnetic Resonance Method to Determine Nuclear Magnetic Moment
19.17	Molecular Orbital Theory of Hydrogen Molecule Ion H_2^+	**19.18**	Heitler-London Theory of Hydrogen Molecule

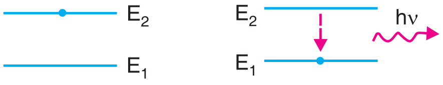
19.1 Induced Absorption, Spontaneous Emission and Stimulated Emission

Three kinds of transition involving electromagnetic radiation are possible between two energy levels, E_1 and E_2, in an atom (Fig. 19.1).

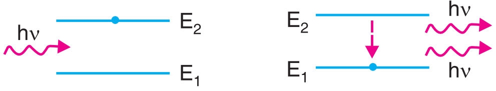
(*i*) If the atom is initially in the lower state E_1, it can be raised to E_2 by absorbing a photon of energy $E_2 - E_1 = h\nu$. This process is called *induced absorption.* Einstein postulated that the *induced*



(a) Induced absorption



(b) Spontaneous emission



(c) Stimulated emission

Fig. 19.1

absorption transition rate was proportional to the number of atoms with electrons in the lower state and to the density of radiation energy incident on these atoms or

$$\left.\frac{dN_{12}}{dt}\right|_I = B_{12}\,\rho_{21}\,N_1 \qquad \text{...(1)}$$

where N_1 = No. of atoms with electrons in the $n = 1$ state

ρ_{21} = density of electromagnetic radiation with energy equal to the energy difference between the two states

B_{12} = Einstein coefficient for induced absorption.

(*ii*) If the atom is initially in the upper state E_2, it can drop to E_1 by emitting a photon of energy $h\nu$; this is *spontaneous emission.* Einstein postulated that the spontaneous emission transition rates were proportional to the number of atoms with electrons in the upper state or

$$\left.\frac{dN_{21}}{dt}\right|_s = A_{21}\,N_2 \qquad \text{...(2)}$$

where A_{21} = Einstein coefficient for spontaneous emission.

(*iii*) Einstein suggested that under certain conditions, it is possible to force an excited atom to emit a photon by another matching photon. According to Einstein, an atom in an excited energy state may, under the influence of the electromagnetic field of a photon of frequency ν incident upon it, jump to a lower energy state, emitting an additional photon of same frequency ν. Thus now two photons instead of one move on. In short, the excited atom emits light waves in step with the incoming wave and increases its intensity. This is known as *stimulated emission* of radiation. The radiated light waves are exactly in phase with the incident ones. So the result is an enhanced beam of coherent light.

The transition rate for stimulated emission is proportional to N_2 and to the density of radiation incident on the atoms with energy equal to the energy difference between the two states or

$$\left.\frac{dN_{21}}{dt}\right|_S = B_{21}\,\rho_{21}\,N_2 \qquad \text{...(3)}$$

where B_{21} = the Einstein coefficient for stimulated emission.

After making the above assumptions, Einstein showed that for thermal equilibrium the coefficients of induced absorption and stimulated emission are equal:

$$B_{21} = B_{12} \qquad \text{..(4)}$$

He also showed that the relationship between the coefficient of spontaneous emission and the coefficient of stimulated emission is

$$\frac{A_{21}}{B_{21}} = \frac{8(E_2 - E_1)}{h^2 c^3} \qquad \text{...(5)}$$

where $E_2 - E_1$ is the energy difference between the two states.

Principle of Laser: Let us consider an assembly of atoms of some kind that have metastable states of excitation energy $h\nu$. Suppose we somehow raise a majority of the atoms to the metastable level. If we now shine light of frequency ν on the assembly, there will be more induced emission from the metastable level than induced absorption by

Laser

the lower level. The result will be an amplification of the original light. This is the concept that underlies the operation of the laser (Light Amplification by Stimulated Emission of Radiation).

Population Inversion: Under ordinary conditions of thermal equilibrium, the number of atoms in the higher energy state is considerably smaller than the number in the lower energy state ($\because$ by Boltzmann law, $N_2/N_1 = e^{-(E_2 - E_1)/KT}$) *i.e.,* $N_2 < N_1$. Hence there is very little stimulated emission compared with absorption. Let, by some means, the atoms be initially excited so that there are more atoms in the higher energy state E_2 than in the lower energy state E_1. We then have $N_2 > N_1$. This is known as *Population inversion.*

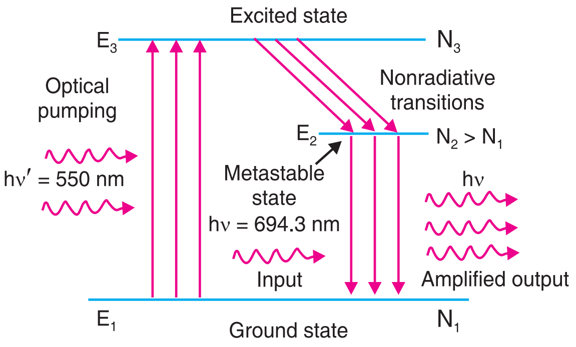

Excited state
E3
N3
Optical pumping
Nonradiative transitions
E2
N2 > N1
hν′ = 550 nm
Metastable state
hν = 694.3 nm
hν
Input
Amplified output
E1
Ground state
N1

Fig. 19.2

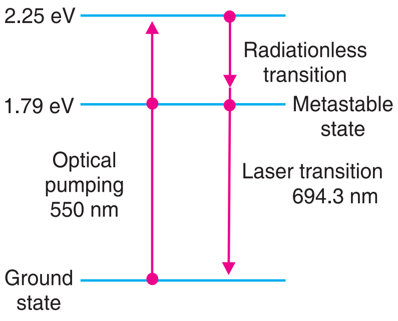

2.25 eV
Radiationless transition
1.79 eV
Metastable state
Optical pumping 550 nm
Laser transition 694.3 nm
Ground state

Fig. 19.3

Pumping: The method of producing population inversion is called *pumping*. One type of pumping is "*optical pumping*". Consider a material whose atoms can reside in three different states as shown in Fig. 19.2. Atoms in ground state are pumped to state E_3 by photons of energy $h\nu' = E_3 - E_1$. The excited atoms then undergo non-radiative transitions with a transfer of energy to the lattice thermal motion, to the level E_2. They remain in this metastable energy state for a comparatively long time. Thus there will be more atoms in the higher metastable energy state E_2 than in the ground state E_1, *i.e.*, we have a "population inversion". Atoms in the metastable state E_2 are now bombarded by photons of energy $h\nu = E_2 - E_1$, resulting in a stimulated emission giving an intense, coherent beam in the direction of the incident photons. This is the method used in the ruby laser (Fig. 19.3).

19.2 The Ruby Laser

It consists of a ruby cylindrical rod whose ends are optically flat and accurately parallel (Fig. 19.4). One end is fully silvered and the other is only partially silvered. The rod is surrounded by a glass tube. The glass tube is surrounded by a helical xenon flash tube which acts as the optical pumping system.

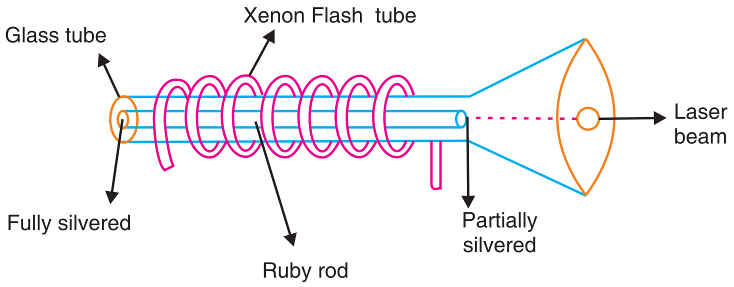

Xenon Flash tube
Glass tube
Laser beam
Fully silvered
Partially silvered
Ruby rod

Fig. 19.4

The ruby rod is a crystal of aluminium oxide (Al_2O_3) doped with 0.05% chromium oxide (Cr_2O_3), so that some of the aluminium atoms in the crystal lattice are replaced by Cr^{3+} ions. The energy level diagram of these chromium ions is shown in Fig. 19.2. The *Cr* ions are excited from level E_1 to level E_3 by the absorption of light of wavelength 550 nm from the xenon flash tube. The excited ions quickly undergo nonradiative transitions with a transfer of energy to the lattice thermal motion, to the level E_2. The E_2 level is a metastable state with a lifetime of about 3×10^{-3} s. (usual atomic lifetimes are $\approx 10^{-8}$ s). Now, the population of the E_2 level becomes greater than that of the E_1 level. Thus "population inversion" is achieved.

Some photons are produced by spontaneous transition from E_2 to E_1, and have a wavelength of 694.3 nm (ruby red). The ends of the ruby rod act as reflecting mirrors. Therefore, photons that are not moving parallel to the ruby rod escape from the side, but those moving parallel to it are reflected back and forth. These stimulate the emission of similar other photons. The chain reaction quickly develops a beam of photons all moving parallel to the rod, which is monochromatic and is coherent. When the beam develops sufficient intensity, it emerges through the partially silvered end.

Ruby Laser.

Once all the chromium ions in the metastable level have returned to ground level, the laser action stops. It is then necessary to send one more flash of pumping radiation through the rod. Thus the ruby laser operates only in pulses.

19.3 Helium-neon Laser

Fig. 19.5 shows the schematic diagram of a He–Ne laser. The laser tube is approximately 5 mm in diameter and 0.5 m long. It contains a helium-neon mixture, in the ratio of 5 : 1, at a total pressure of about 1 torr. The tube has parallel mirrors, one of them partly transparent, at both ends. The spacing of the mirrors is equal to an integral number of half-wavelengths of the laser light. The *He-Ne* mixture contained in the gas discharge tube is ionized by passing a dc current through the gas. Helium atoms are excited very efficiently by electron impact into the 2*s* level (Fig. 19.6) while the neon atoms are much less readily excited by the electrons. This excited 2*s* state of helium is relatively long lived. The energy of this level (20.61 eV) is almost the same as the energy of the 5*s* level (20.66 eV) in neon. Hence the energy of the helium atoms is easily transferred to the neon atoms when they

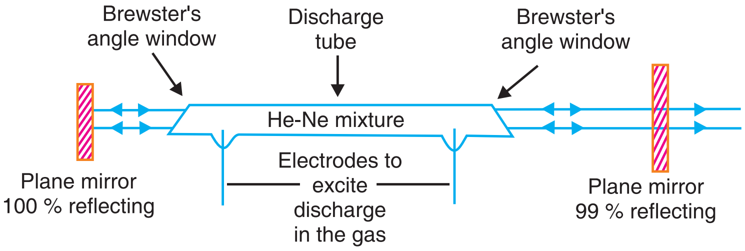

Brewster's angle window
Discharge tube
Brewster's angle window
He-Ne mixture
Plane mirror 100 % reflecting
Electrodes to excite discharge in the gas
Plane mirror 99 % reflecting

Fig. 19.5

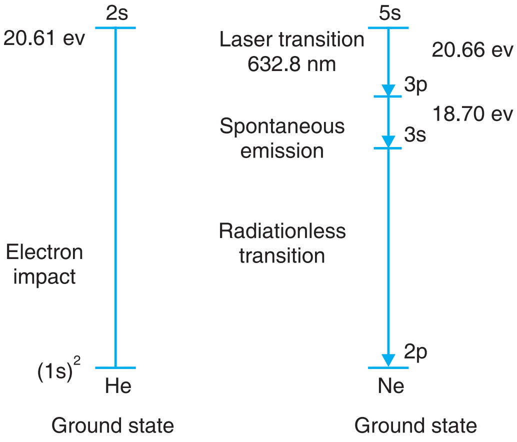
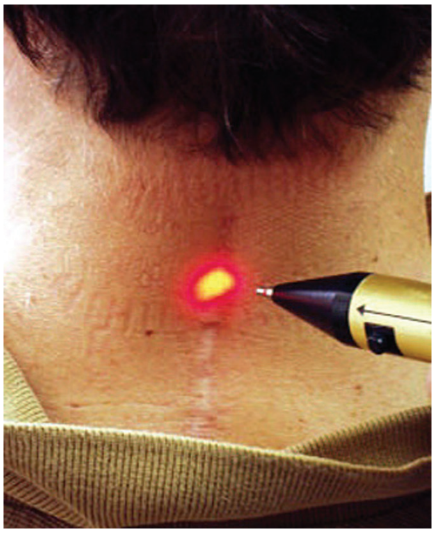

2s
20.61 ev
Electron impact
(1s)²
He
Ground state
Laser transition 632.8 nm
5s
20.66 ev
3p
18.70 ev
3s
Spontaneous emission
Radiationless transition
2p
Ne
Ground state

Fig. 19.6

collide. This preferential transfer of the neon atoms to the 5*s* state results in a population inversion between the 5*s* and 3*p* states. The purpose of the *He* atoms is thus to help achieve a population inversion in the *Ne* atoms. The spontaneous transitions from the 5*s* state to the 3*p* state, produce photons of wavelength 632.8nm, which then trigger stimulated transitions. Photons travelling parallel to the tube are reflected back and forth between the mirrors placed at the ends, and rapidly build up into an intense beam which escapes through the end with the lower reflectivity. The Brewster end windows allow light of one polarisation to pass through without any reflection losses. Because the electron impacts that excite the *He* and *Ne* atoms occur all the time, a *He-Ne* laser operates continuously.



Soft Helium-neon Laser used to Treat the Scar Tissue on the Posterior Neck.

19.4 Semiconductor Laser

The energy emitted from visible LEDs and IR emitters is spontaneous or noncoherent. If the level of current flowing through a GaAs *p-n* junction is increased beyond a certain threshold, a *lasing* action takes place in which the spontaneous emission *stimulates* an increase in the radiant power.

Fig. 19.7 shows the basic structure of a *p–n* junction laser. A pair of parallel planes are cleaved or polished perpendicular to the plane of the junction. The two remaining sides of the diode are roughened to eliminate lasing in directions other than the main one. The structure is called a Fabry-Perot cavity. When a forward bias is applied to the laser diode, a current flows. Initially at low current, there is spontaneous emission in all directions. As the bias is increased, eventually a threshold current is reached at which the stimulated emission occurs and a monochromatic and highly directional beam of light is emitted from the junction.

The advantages of semiconductor lasers are that they are compact, efficient and can be fabricated with ease. However, their monochromaticity, coherence and directionality are inferior to those of other lasers.

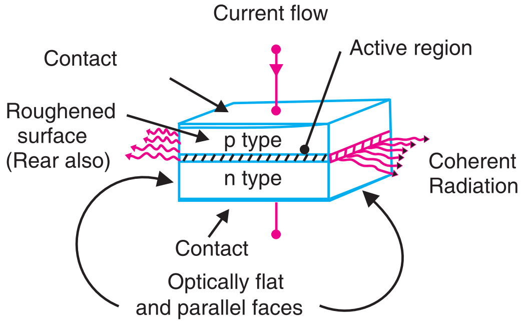
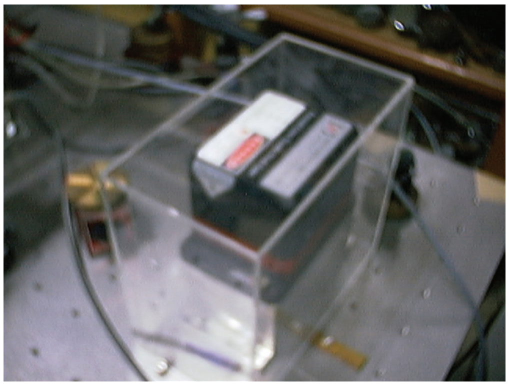


Fig. 19.7



Semiconductor Laser.

19.5 Properties of a Laser Beam

The characteristic properties of a laser beam are (*i*) *monochromaticity* (*ii*) *directionality* (*iii*) *spatial coherence* (*iv*) *temporal coherence* and *brightness.*

(*i*) The light is very nearly monochromatic.

(*ii*) A laser beam diverges hardly at all.

(*iii*) The light is coherent with the waves all exactly in phase with one another.

(*iv*) The beam is extremely intense.

Applications: The laser is used in holography, radio communication in outer space, piercing holes in metals, detecting and ranging objects at great distances, welding, surveying and delicate surgery.

Holography: Holography is a method of producing optical images with full three-dimensional perspective, in contrast to conventional photography that records only a two-dimensional projection. The problem with photography is that only the intensity of the light scattered from an object is recorded, whereas for a full representation we require both the amplitude and the phase. We can record the phase information by mixing the light scattered from the laser beam with a coherent reference beam. This is illustrated in Fig. 19.8. Let us suppose that the reference amplitude is a maximum (+1) at the photographic plate at *A*, *B* and *C*, but the relative phases of the scattered beams are such as to give 0, –1, and +1 at *A,B* and *C,* respectively.

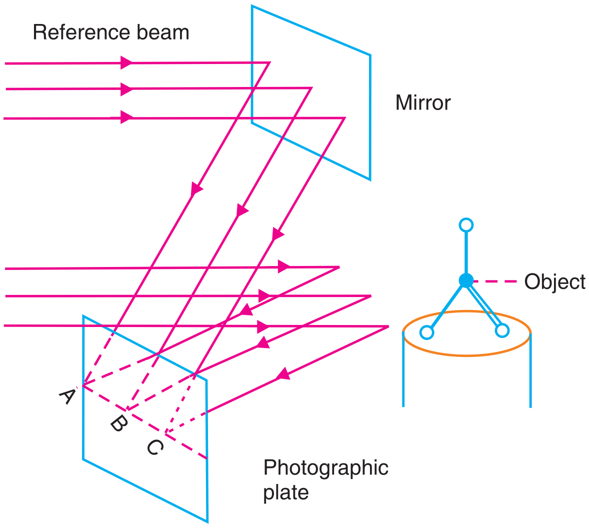


Fig. 19.8

Then, if the light is coherent, we must add amplitudes, before squaring to give the intensity. This gives resultant intensities of 1, 0 and 4 at *A,B* and *C*. But in the absence of the reference beam the intensities would be 0, 1, and 1. We have therefore recorded on the photographic plate the information that the phase of the scattered light at *B* differs from that at *C*. Clearly, for this process to work, the light must be coherent over a reasonable area. If the photographic plate is developed and illuminated with a coherent beam similar to that used as a reference, an observer sees a complete, three-dimensional reconstruction of the object. As he moves his head, the perspective of the image changes in exactly the same way as that of the object.

MOLECULAR SPECTRA

19.6 Introduction

Apart from the ground state, a molecule can be in a higher energy state, and transitions between the various energy levels give rise to the observed molecular spectra.

The energy of a diatomic molecule arises from three modes:

(*i*) the electronic configuration of the electrons in the molecule,

(*ii*) the vibration of the atoms about the equilibrium position, and

(*iii*) the rotation of the molecule as a whole about its centre of mass.

The total energy of a molecule can be expressed as the sum of three independent terms:

$$E = E_e + E_v + E_r \qquad \text{...(1)}$$

$$E_e,\ E_v,\ E_r \text{ are quantised and } \qquad E_v > E_r.$$

If the molecule remains in its ground state level of electronic and vibrational energies, and if it suffers transitions between different rotational energies, we get the *pure rotational spectrum* of the molecule. These spectra are in the microwave and far infrared regions.

If the molecule remains in its ground state level of electronic energy, and if transitions occur between different vibrational and rotational energies we get the *vibration-rotation* spectrum of the molecule. These spectra are in the near infrared region.

If transitions occur between different electronic-vibrational-rotational energy levels, we get the *electronic spectrum* of the molecule. These spectra are in the visible and ultraviolet regions.

19.7 Theory of the Origin of Pure Rotational Spectrum of a Molecule

We may picture a diatomic molecule as consisting of atoms of masses m_1 and m_2 a distance R apart. The diatomic molecule can rotate about its centre of mass as shown in Fig. 19.9.

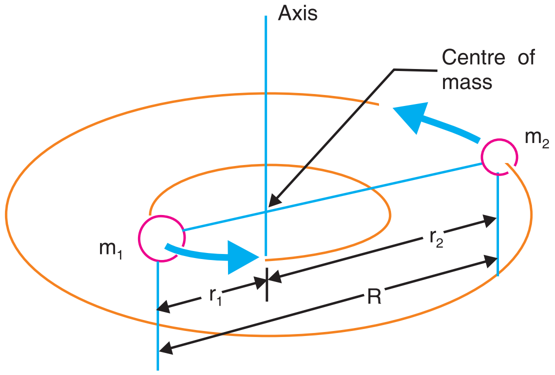


Fig. 19.9

The moment of inertia of this molecule about an axis passing through its centre of mass and perpendicular to a line joining the atoms is

$$I = m_1 r_1^2 + m_2 r_2^2 \qquad \text{...(1)}$$

where r_1 and r_2 are the distances of m_1 and m_2, respectively from the centre of mass.

From the definition of centre of mass

$$m_1 r_1 = m_2 r_2 \qquad \text{...(2)}$$

Further $$R = r_1 + r_2 \qquad \text{...(3)}$$

From equations (2) and (3)

$$r_1 = \frac{m_2}{m_1 + m_2} R \text{ and } r_2 = \frac{m_1}{m_1 + m_2} R.$$

Putting these values of r_1 and r_2 in Eq. (1),

$$I = \frac{m_1 m_2^2}{(m_1 + m_2)^2} R^2 + \frac{m_2 m_1^2}{(m_1 + m_2)^2} R^2 = \frac{m_1 m_2}{(m_1 + m_2)} R^2$$

or $$I = m' R^2 \qquad \text{...(4)}$$

where the reduced mass m' is defined as

$$m' = \frac{m_1 m_2}{(m_1 + m_2)} \qquad ...(5)$$

The angular momentum L of the molecule has the magnitude $L = I\omega$.

The angular momentum of the rotating diatomic molecule is quantised according to

$$L = \sqrt{J(J+1)}\,\hbar \quad J = 0, 1, 2, 3, ... \qquad ...(6)$$

J is called the rotational quantum number.

The rotational kinetic energy of a diatomic molecule is $\frac{1}{2} I\omega^2$.

$$\therefore \qquad E_J = \frac{1}{2} I\omega^2 = \frac{L^2}{2I} = \frac{J(J+1)\hbar^2}{2I} \qquad ...(7)$$

Rotational transitions occur only in those molecules which possess a permanent electric dipole moment. For this reason nonpolar diatomic molecules such as H_2 and symmetric polyatomic molecules such as CO_2 and CH_4 do not exhibit rotational spectra.

The selection rule for rotational transitions is $\Delta J = \pm 1$ ...(8)

Rotational spectra are always obtained in absorption. Hence for a transition from some initial state of quantum number J to the next higher state of quantum number $J + 1$, the frequency of the absorbed photon is

$$\nu_{j \to J+1} = \frac{\Delta E}{h} = \frac{E_{J+1} - E_J}{h} = \frac{\hbar}{2\pi I}(J+1) \qquad ...(9)$$

The spectrum of a rigid molecule therefore consists of equally spaced lines (Fig. 19.10).

Importance of the pure rotational spectrum for determining the properties of a molecule: The frequency difference between two successive lines in the pure rotational spectrum of a diatomic molecule is given by

$$\Delta\nu = \nu_{J+1 \to J+2} - \nu_{J \to J+1} = \frac{\hbar}{2\pi I}.$$

By measuring $\Delta\nu$, the moment of inertia I of the molecule can be calculated. Now,

$$I = m' R^2 = \left\{ \frac{m_1 m_2}{(m_1 + m_2)} \right\} R^2$$

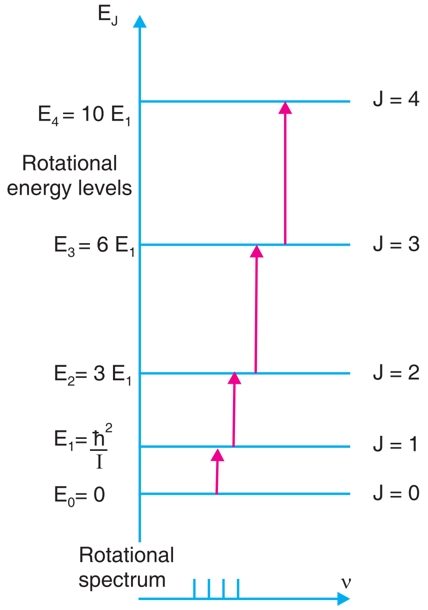


Fig. 19.10

Thus the distance R between the nuclei can be calculated. This gives the length of the chemical bond between the atoms.

Example: *In the CO molecule the wavenumber difference between the successive absorption lines in the pure rotational spectrum is 384* m^{-1}. *Calculate the moment of inertia of the molecule and the equilibrium bond length of the molecule. Masses of the* C^{12} and O^{16} *atoms are respectively* 1.99×10^{-26} *kg and* 2.66×10^{-26} *kg.*

Sol. Here, $\Delta\bar{\nu} = \frac{\hbar}{2\pi I c} = 384\, m^{-1}$

$$\therefore \qquad I = \frac{\hbar}{2\pi c\Delta\bar{\nu}} = \frac{1.054\times10^{-34}}{2\pi\times(3\times10^{8})\times384} = 1.456\times10^{-46}\ kg\, m^2$$

The reduced mass *m'* of the *CO* molecule is

$$m' = \frac{m_1\, m_2}{m_1+m_2} = \left[\frac{1.99\times2.66}{1.99+2.66}\right]\times10^{-26}\ \text{kg} = 1.138\times10^{-26}\ \text{kg}.$$

$$\therefore \qquad \text{Bond length} = R = \sqrt{\frac{I}{m'}} = \sqrt{\frac{1.456\times10^{-46}}{1.138\times10^{-26}}} = 0.1131\ \text{nm}\ .$$

19.8 Theory of the Origin of the Vibration-rotation Spectrum of a Molecule

When sufficiently excited, a molecule can vibrate as well as rotate. We shall consider only diatomic molecules. Fig. 19.11 shows how the potential energy of a diatomic molecule varies with the internuclear distance *R*. In the neighbourhood of the minimum of this curve, which corresponds to the normal configuration of the molecule, the shape of the curve is very nearly a parabola. In this region, then

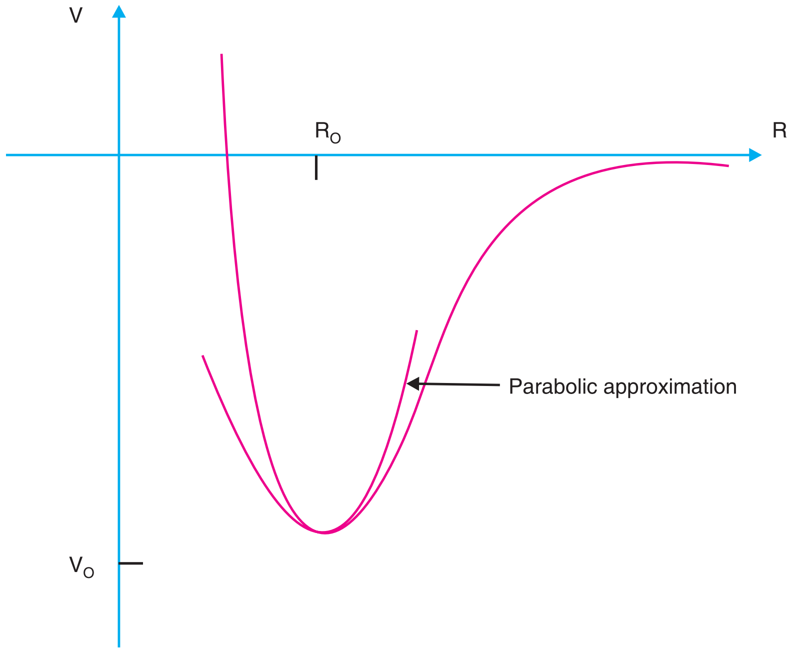


Fig. 19.11

$$V = V_0 + \frac{1}{2}k(R-R_0)^2 \qquad ...(1)$$

where R_o is the equilibrium separation of the atoms.

The interatomic force that gives rise to this potential energy

$$F = -dV/dR = -k\,(R-R_o) \qquad ...(2)$$

The force is just the restoring force that a stretched or compressed spring exerts and, as with a spring, a molecule suitably excited can undergo simple harmonic oscillations. Classically, the frequency of a vibrating body of mass *m* connected to a spring of force constant *k* is

$$\nu_0 = \frac{1}{2\pi}\sqrt{\frac{k}{m}} \qquad \text{...(3)}$$

In the diatomic molecule we have two bodies of masses m_1 and m_2 joined by a spring. The use of the “reduced mass” reduces the problem of two particles m_1 and m_2 vibrating about their common centre of mass to the problem of a single mass m' vibrating about a fixed point (Fig. 19.12).

The frequency of oscillation of such a two-body oscillator is

$$\nu_0 = \frac{1}{2\pi}\sqrt{\frac{k}{m'}} \qquad \text{...(4)}$$

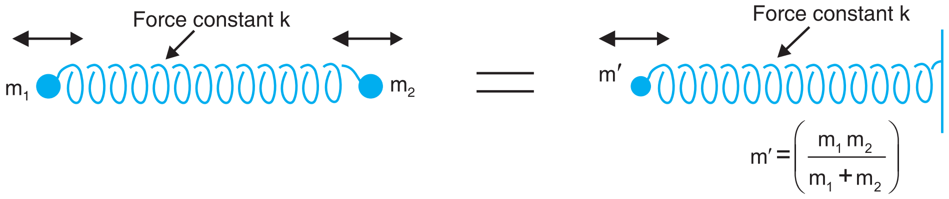


Fig. 19.12

When the harmonic oscillator problem is solved quantum mechanically, the energy of the oscillator is found to be restricted to the values

$$E_v = (v+\frac{1}{2})\, h\nu_0 \qquad \text{...(5)}$$

Here v is the vibrational quantum number. It may have the values $v = 0, 1, 2, 3, ...$

Substituting the value of ν_o from Eq. (4),

$$E_v = (v+\frac{1}{2})\hbar\sqrt{\frac{k}{m'}} \qquad \text{...(6)}$$

These levels are sketched on Fig. 19.13.

For *large*-amplitude vibrations, the parabola does not fit the actual energy curve. Eq. (6) no longer holds. The energy well is wider than the parabola for the higher E_v’s, and therefore these energy levels are closer together than the lower E_v’s. This diagram also shows the fine structure in the vibrational levels caused by the simultaneous excitation of rotational levels.

The selection rule for transitions between vibrational states is

$$\Delta v = \pm 1 \qquad \text{...(7)}$$

Pure vibrational spectra are observed only in liquids. This is because interactions between neighbouring molecules prevent their rotational motion.

The vibrations and rotations of a molecule take place independently of each other. Under these circumstances the energy levels of a diatomic molecule are specified by

$$E_{v,J} = (v+\frac{1}{2})\,\hbar\sqrt{\frac{k}{m'}} + J(J+1)\,\frac{\hbar^2}{2I} \qquad \text{...(8)}$$

E_v is very much larger than E_r. Even at high temperatures only the vibrational states corresponding to $v = 0$ and $v = 1$ are excited. Fig. 19.14 shows the $J = 0, 1, 2, 3$ and 4 levels of a diatomic molecule for the $v = 0$ and $v = 1$ vibrational states. It also shows the spectral lines in absorption. The selection rules are $\Delta v = 1$ and $\Delta J = \pm 1$. The transition with $\Delta J = 0$ is not allowed.

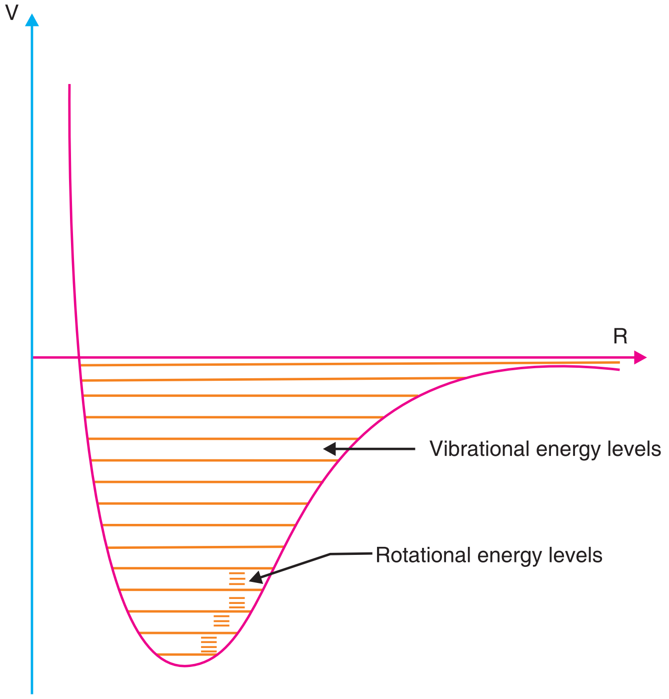


Fig. 19.13

The $v = 0 \longrightarrow v = 1$ transitions fall into two categories :

(*i*) *P Branch:* Here, $\Delta J = -1$ (*i.e.*, $J \rightarrow J - 1$).

From Eq. (8) the frequencies of the spectral lines in the *P* branch are given by

$$v_p = \frac{E_{1,J-1} - E_{0,J}}{h} = \frac{1}{2\pi}\sqrt{\frac{k}{m'}} + [(J-1)\,J - J\,(J+1)]\,\frac{\hbar}{4\pi I}$$

or

$$v_p = v_0 - J\,\frac{\hbar}{2\pi I} \qquad J = 1, 2, 3, \ldots \qquad \ldots(9)$$

(*ii*) *R Branch*: Here, $\Delta J = +1$ (*i.e.*, $J \rightarrow J + 1$)

$$v_R = \frac{E_{1,J+1} - E_{0,J}}{h} = \frac{1}{2\pi}\sqrt{\frac{k}{m'}} + [(J+1)\,(J+2) - J\,(J+1)]\,\frac{\hbar}{4\pi I}$$

or

$$v_R = v_0 + (J+1)\,\frac{\hbar}{2\pi I} \qquad J = 0, 1, 2, 3, \ldots \qquad \ldots(10)$$

There is no line at $v = v_0$ (the *Q* branch) because transitions for which $\Delta J = 0$ are forbidden in diatomic molecules.

The spacing between the lines in both the *P* and the *R* branch is $\Delta v = \dfrac{\hbar}{2\pi I}$.

Hence, by measuring the frequencies of these lines, the moment of inertia of the molecule and the length of the chemical bond can be calculated.

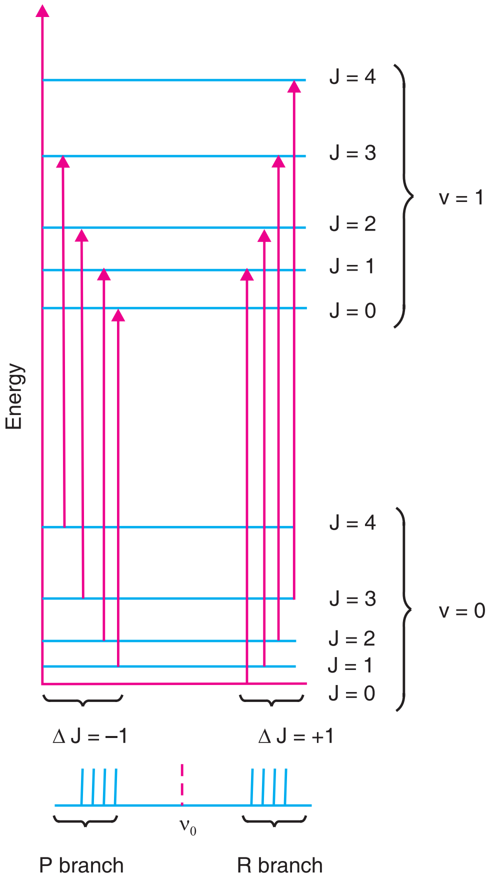

Energy
J = 4
J = 3
J = 2
J = 1
J = 0
v = 1
J = 4
J = 3
J = 2
J = 1
J = 0
v = 0
Δ J = –1
Δ J = +1
ν_0
P branch
R branch

Fig. 19.14

Fig. 19.15 shows a typical vibration-rotation molecular spectrum of *HBr*. It is made by sending radiation from a source of continuous infrared wavelengths through a cell containing *HBr* and recording the absorption spectra.

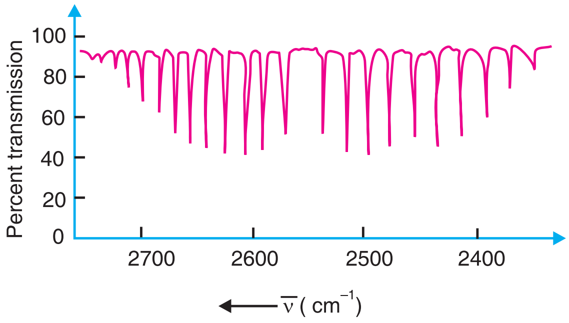


Fig. 19.15

19.9 Electronic Spectra of Molecules

When there is enough available energy, the electronic states of the molecules can be excited. This gives rise to the most general type of transitions, ***rotation-vibration-electronic transitions.*** These produce photons of frequency

$$\nu = \frac{E_{e1} - E_{e2}}{h} + \frac{E_{v1} - E_{v2}}{h} + \frac{E_{rot1} - E_{rot2}}{h}.$$

It is the outermost electrons of the atoms composing the molecules that are involved. These transitions are energetic enough that they produce *electronic bands* is the visible and ultraviolet regions. All molecules exhibit electronic spectra, since a dipole moment change always accompanies a change in the electronic configuration of a molecule. In molecular spectra, as in atomic spectra, radiative transitions between electronic states of different total spin are forbidden.

19.10 Rayleigh's Scattering

Rayleigh showed theoretically that the intensity of scattered light is inversely proportional to the fourth power of the wavelength *i.e.*, $I \propto \frac{1}{\lambda^4}$. This is known as Rayleigh's law of scattering. I is also proportional to the square of the volume of the particle. The blue colour of the sky is due to the greater scattering of short wavelengths by the dust suspensions in the atmosphere or by the air molecules. This type of scattering simply produces separation of wavelengths originally present in the incident light. No wavelength can be present in the scattered light which is not present in the incident light *i.e.*, no new wavelength is created. For this reason, Rayleigh scattering is called coherent scattering.

RAMAN EFFECT

19.11 Discovery

While studying the scattering of light, Raman found that when a beam of monochromatic light was passed through organic liquids such as benzene, toluene, etc., the scattered light contained other frequencies in addition to that of the incident light. This is known as *Raman effect.*

The original arrangement used by Raman was simple in design. A round-bottomed glass flask was filled with pure dust-free benzene. The liquid was strongly illuminated by the mercury blue line of wavelength 4358 Å. The scattered light was examined through a spectroscope in a direction perpendicular to that of the incident beam. It was observed that the spectrum contained, in addition to the original wavelength (4358 Å), some lines which had wavelengths longer as well as shorter than the original wavelength. These lines of modified wavelengths are known as *Raman lines.* The lines having wavelengths greater than that of the incident wavelength are called *Stokes lines* and those having shorter wavelengths are called *anti-Stokes lines.* The stokes lines are found to be more intense than the anti-Stokes lines. Most of the Raman lines are strongly polarised.

C.V. Raman (1888-1970)

The displacement of the modified spectral lines from the exciting line, when measured in wavenumbers, depends only on the scattering substance and is independent of the wavenumber of the exciting radiation. The Raman spectrum is thus characteristic of the scattering substance.

Raman effect is quite different from Rayleigh scattering. In Rayleigh or coherent scattering, there is no change in wavelength. But in Raman scattering, the scattered light contains modified wavelengths as well as the original wavelengths. For this reason, Raman effect is also called *incoherent scattering* and is considered as the *optical analogue of the Compton effect.*

19.12 Experimental Study of Raman Effect

The apparatus shown in Fig. 19.16 first developed by Wood, is used for studying Raman effect in liquids. It consists of a glass tube *AB* containing the pure experimental liquid free from dust and air bubbles. The tube is closed at one end by an optically plane glass plate *W* and at the other end it is drawn into a horn (*H*) and blackened on the outside. Light from a mercury arc *S* is passed through a filter *F* which allows only monochromatic radiation of λ = 4358 Å to pass through it. The tube is surrounded by a water-jacket (*J*) through which water is circulated to prevent overheating of the liquid. A semi-cylindrical aluminium reflector *R* is used to increase the intensity of illumination. The scattered light coming out of *W* is condensed on the slit of a spectrograph. The spectrograph must have a large light gathering power and the prism must have a large resolving power. A short focus camera is used to photograph the spectrum.

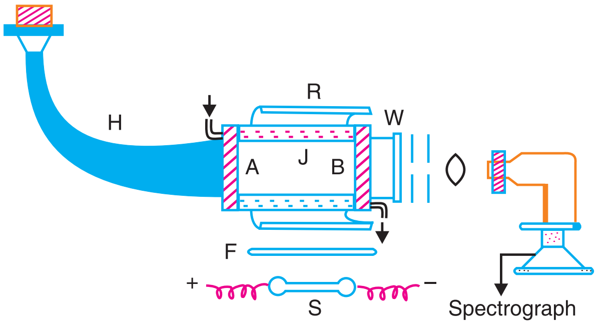


Fig. 19.16

Experimental Study of Raman Effect.

On developing the photographic plate, it exhibits a number of Stokes' lines, a few anti-Stokes lines and a strong unmodified line (Fig. 19.17).

Characteristics of Raman Lines: (1) The Stokes lines are always more intense than antistokes lines.

(2) The Raman lines are symmetrically displaced about the parent line.

(3) The frequency difference between the modified and parent line represents the frequency of the corresponding infrared absorption line.

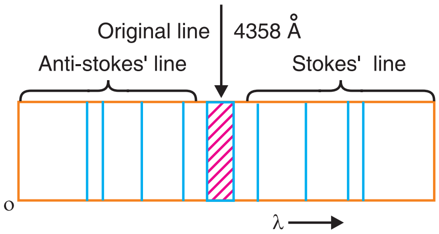


Fig. 19.17

19.13 Quantum Theory of Raman Effect

Raman effect is due to the interaction between a light photon and a molecule of the scatterer. Suppose a photon of frequency ν_1 is incident on a molecule and there is a collision between the two. Let m = mass of the molecule, v_1 and v_2 its velocities before and after impact, E_1 and E_2 the intrinsic energies of the molecule before and after collision. Let ν_2 be the frequency of the scattered photon.

Applying the principle of conservation of energy,

$$E_2 + \frac{1}{2}mv_2^2 + h\nu_2 = E_1 + \frac{1}{2}mv_1^2 + h\nu_1 \qquad ...(1)$$

We may assume that the K.E. of the molecule is unaltered during the process. Hence,

$$E_2 + h\nu_2 = E_1 + h\nu_1 \text{ or } \nu_2 - \nu_1 = \frac{E_1 - E_2}{h}$$

or

$$\nu_2 = \nu_1 + \frac{E_1 - E_2}{h} \qquad ...(2)$$

Three cases may arise:

(1) When there is no change in the energy of the molecule, $E_1 = E_2$. Then $\nu_1 = \nu_2$. This represents the unmodified line.

(2) If $E_2 > E_1$, then, $\nu_2 < \nu_1$. This represents the Stokes line. It means that the molecule has absorbed some energy from the incident photon. Consequently the scattered photon has lower energy or longer wavelength.

(3) If $E_2 < E_1$, then, $\nu_2 > \nu_1$. This represents the anti Stokes line. It means that the molecule was previously in the excited state and it handed over some of its intrinsic energy to the incident photon. The scattered photon thus has greater energy or shorter wavelength.

Since the molecules possess quantised energy levels, we can write,

$$E_1 - E_2 = nh\nu_c \qquad ...(3)$$

where $n = 1, 2, 3....$etc, and ν_c = the characteristic frequency of the molecule.

In the simplest case $n = 1$, equation (2) reduces to

$$\nu_2 = \nu_1 \pm \nu_c \qquad ...(4)$$

Equation (4) shows that the frequency difference $\nu_1 - \nu_2$ between the incident and scattered photon corresponds to the characteristic frequency ν_c of the molecule.

Raman Effect and Fluorescence: Raman effect and fluorescence resemble each other in that the incident light in both the cases suffers a change of wavelength and lines of new wavelengths appear in the spectrum of either. But the Raman effect is quite different from fluorescence. We compare the two phenomena in the table.

Raman Spectra	*Fluorescence Spectra*
1. Spectral lines have frequencies greater and lesser than the incident frequency.	The frequencies of the lines in the fluorescent spectrum are always less than the incident frequency.
2. The frequency shifts of the Raman lines are determined by the scatterer rather than the frequencies themselves.	Frequencies of the flourescent lines are determined by the nature of the scatterer.
3. Raman lines are strongly polarised.	Lines are not polarised.

19.14 Applications

(1) Raman Effect and Molecular Structure: A qualitative identification of Raman spectrum consists in the evaluation of the wavelength of the lines, their intensities, and state of polarisation. Investigation of bond angles, bond stiffness, and other structural confirmation require Raman data in addition to infrared studies.

***(i)* Diatomic Molecules:** From the analysis of Raman spectra of a diatomic molecule, we can have an idea about the nature of the chemical bond existing between the atoms. In a diatomic molecule, the frequency of vibration of the atoms is given by $\nu = \frac{1}{2\pi}\sqrt{\frac{F}{\mu}}$ where F is the restoring force per unit displacement and μ is the reduced mass of the molecule. It is seen that a molecule in which the force binding the atoms is great should have higher characteristic frequency than one in which the force is weak. This force depends upon the nature of interatomic bonds. In covalent molecules polarisability is considerably changed by the nuclear oscillations due to the nuclei. This appreciable change in polarisability gives rise to intense Raman lines. In electrovalent molecules, the binding elecrons definitely change over from one nucleus to the other in the formation of the molecule so that the polarisability of the molecule is little affected by nuclear oscillations and hence no Raman lines will appear.

(*ii*) Triatomic molecules: Dealing with triatomic molecules (of type AB_2), the questions to be decided are: whether each molecule is linear or not and, if linear, whether it is symmetrical (*B-A-B*) or asymmetrical (*B-B-A*). From the number and intensity of the observed lines in the Raman effect in conjunction with infrared data, it is possible to draw important conclusions about molecular structure. Theory leads to the following rule, known as the *rule of mutual exclusion.* It states that *for molecules with a centre of symmetry transitions that are allowed in the infrared are forbidden in Raman spectra and vice versa.* The rule does not imply that all transitions forbidden in one must occur in the other; *i.e.*, some transitions may be forbidden in both. On the other hand certain transitions can occur both in the infra-red and in the Raman spectra in the case of molecules without a centre of symmetry.

Examples: (*a*) CO_2 has two very strong bands in its infrared absorption spectrum at 66800 and 234900 m^{-1} while only one strong band in its Raman spectrum at 138900 m^{-1}. None of these bands occur both in Raman and infrared spectra. Hence, it follows from the rule of mutual exclusion, that CO_2 molecule has a centre of symmetry. This implies that the molecule is linear and symmetric and hence it should be represented by $O - C - O$.

(*b*) Nitrous oxide (N_2O) has three absorption bands at 222400, 128500 and 58900 m^{-1} of which the first two appear in the Raman spectrum. Thus the molecule cannot have a centre of symmetry though linear. Hence the molecule has the unsymmetrical structure $N - N - O$.

In a similar manner, the bent symmetric structure of a water molecule represented by Fig. 19.18 is revealed by Raman Effect. Thus the study of Raman spectra of different substances enables one to classify them according to their molecular structure.

(2) Raman effect in crystals is complementary to X-ray crystal study and provides information about the binding forces in crystals.

(3) Various chemical effects like strength of chemical bonds, electrolytic dissociation, hydrolysis, etc., have been understood through Raman effect. Specific heat capacities of solids, brilliance of metals and their molecular structure have been explained by Raman effect.

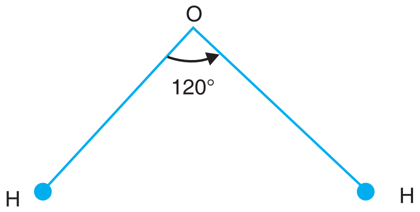


Fig. 19.18

19.15 Laser Raman Spectroscopy

Raman spectroscopy is essentially emission spectroscopy. In Raman spectroscopy the exciting source plays a predominant part. Before the discovery of the laser, the exciting source was invariably a mercury lamp. Fig. 19.19 shows excitation of a sample

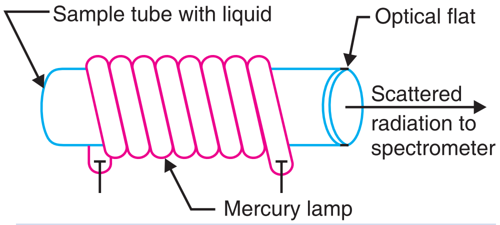


Fig. 19.19

with a spiral mercury lamp. This form of excitation pumps a good deal of energy into the sample. But it is spread over the many emission lines of the mercury vapour such as 435.8 and 253.6 nm. This excitation technique has three disadvantages.

1. The extended source allows a good deal of the exciting radiation to be scattered directly into the spectrometer. As a result, the Raman lines are masked.

2. The sample tube needs to be about 20–30 cm long and 1–2 cm in diameter, so that a considerable quantity of sample is required.

3. The relatively high frequency of the exciting radiation causes flourescence in the sample. The resulting flourescence spectrum swamps the very weak Raman spectrum.

Before 1966 Raman spectroscopy was not used for routine analytical work, as was the case with IR. The reason for this was that the scattered radiation was very weak in intensity. It took several hours or even days to record a spectrum. Further, the process of obtaining a good Raman spectrum required 10–20 ml of sample. However, with the use of laser as the exciting source, the scope of Raman spectroscopy is greatly widened. Now Raman spectra of all samples can be run on a completely routine basis using one mililitre or less of sample and taking a few minutes only.

Laser Raman Spectroscopy

The advantages of using laser are:

1. A single, intense frequency source replaces the multiple-lined mercury lamp. No filtering is necessary.

2. The resolution is better since the linewidth of a laser line is smaller than that of a mercury-excited line.

3. Laser light is highly coherent. Hence it can be focussed finely into a small sample. Thus a relatively large power is packed in a small space.

4. A large choice of exciting frequencies are available, and it is possible to study coloured solutions without causing any electronic transitions. This is of great practical importance in studying inorganic salts in solution since many of them are coloured compounds. The use of IR is ruled out since the solvent is usually water, which absorbs strongly in the IR region.

Fig. 19.20 shows how a laser beam may be used to irradiate a sample using a number of mirrors. By careful alignment as many as 10 reflections can be arranged. As a result, we get a highly enhanced Raman signal. Only a small quantity of the sample is required. It may be enclosed in a quartz capillary tube and the laser beam directed along its length. Solids in powder form or in the form of transparent blocks can be used. Lasers operate at frequencies lower than that of a mercury lamp. Hence there is less chance of fluorescence.

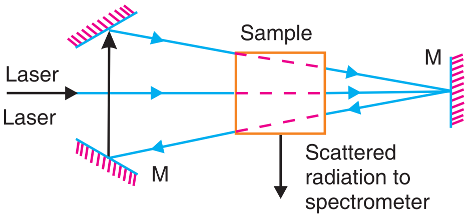


Fig. 19.20

Radiation scattered from the sample is directed, via mirrors, into a spectrometer operating in the visible region. The monochromator is either a quartz prism or a grating. The radiation is detected by a photoelectric detector with its output fed to an amplifier and pen recorder.

19.16 Rabi's Molecular Beam Magnetic Resonance Method to Determine Nuclear Magnetic Moment

Molecular beam apparatus used by Rabi for nuclear magnetic resonance experiments is shown in Fig. 19.21. *O* is the oven where molecules of LiCl are produced in the form of a beam which is suitably collimated by the slits S_1 and S_2. The beam then passes through the magnets *A, C* and *A'*. *A* and *A'* produce strong inhomogeneous identical magnetic fields. The vertical magnetic field gradient in *A'* is opposite to that in *A* (Fig. 19.21). The magnet *C* provides a uniform vertical magnetic field B_0. The hair pin loop *R* placed between the poles of *C* provides an oscillating horizontal magnetic field.

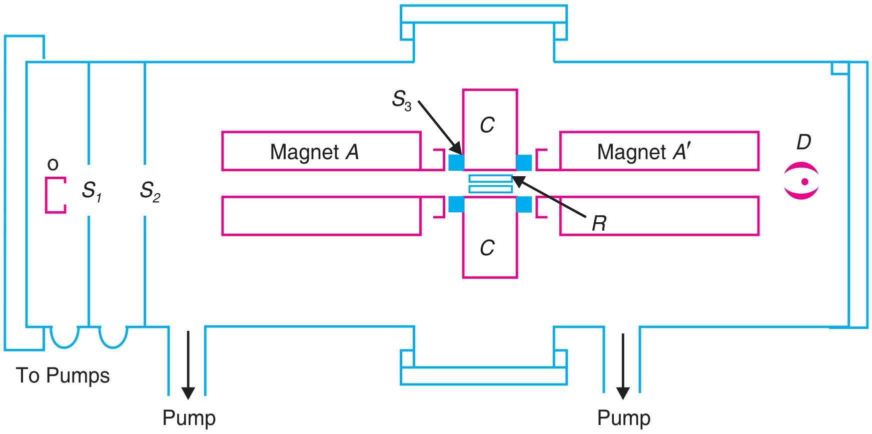
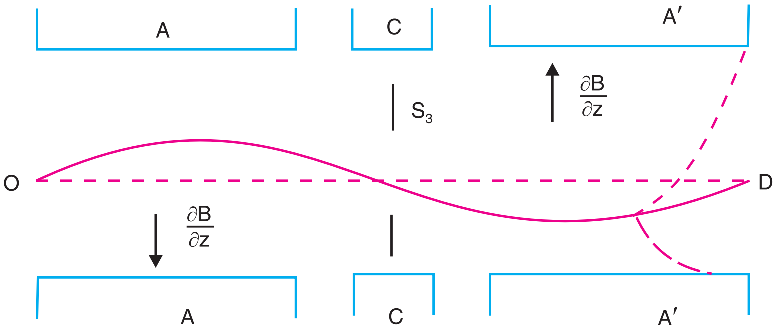


Fig. 19.21

Let the magnetic fields be parallel to *z*-direction. A translational force F_z acts on the molecule parallel to *z*-direction due to field gradient $\dfrac{\partial B}{\partial z}$.

$$F_z = \frac{\mu_I . M_I}{I} \cdot \left(\frac{\partial B}{\partial z}\right) \qquad ...(1)$$

where $\dfrac{\mu_I M_I}{I}$ is the magenetic moment of different sub-states.



A
C
A′
S_3
$\frac{\partial B}{\partial z}$
O
D
$\frac{\partial B}{\partial z}$
A
C
A′

Fig. 19.22

The molecules with Maxwellian distribution of velocity are moving with an average speed of *v* in the magnet *A*. If the length of this magnet is *L*, the molecules with speed *v* spend a time *L* /*v* and its vertical displacement when leaving this field is given by

$$Z_0 = \frac{\mu_I M_I (\partial B/\partial z) L^2}{4E} \qquad \text{...(2)}$$

where E is the energy of the molecule. Molecules with various magnetic quantum states M_I which are making various inclinations to the horizontal axis will be passing through S_3. The magnetic field under A' and the gradient $\partial B/\partial z$ are such that the force is opposite to that produced by A. Thus the molecules which have passed through S_3 will be able to be collected at D. The detector D is a Wolfram wire (t = 1000ºC) which has large work function than Li alkali atoms. When Li neutral atom falls on the hot wire, it gives its electron to the wire and the positive ion is boiled off. The detector electron current measures the intensity of the ion beam. If the orientation of the molecules changes, then they will not be able to reach the detector D and the intensity recorded will be reduced. In Fig. 19.23, the solid line shows the path of a molecule whose spin remains unchanged. The dashed lines show the paths of molecules whose spins change orientation while passing through the magnet C.

There will be strong resonance corresponding to the change in M_I given by the selection rule $\Delta M_I = \pm 1$. The resonance frequency is related to the magnetic moment and the steady field as

$$h\nu = g_I \mu_N B = \frac{\mu_I}{I} . \mathrm{B}$$

In actual experiment, ν is kept constant and resonance is obtained by changing B_0. B_0 is the field under magnet C which is also directed upwards.

If the magnetic dipole moment μ_I is aligned along the magnetic field and is taken to a direction opposite to the magnetic field, the total work done will be

$$E = 2\,\mu_I . B.$$

In this transition there will be $(2I + 1)$ magnetic sub-levels and $2I$ level intervals between the adjacents. Hence the energy difference between the adjacent levels will be given by $\frac{2\mu_I . B}{2I} = \frac{\mu_I B}{I}$ which will be equal to the energy of resonance photon $h\nu$. Rabi obtained following curves for resonance by changing the field B_0 and keeping ν constant (Fig. 19.23).

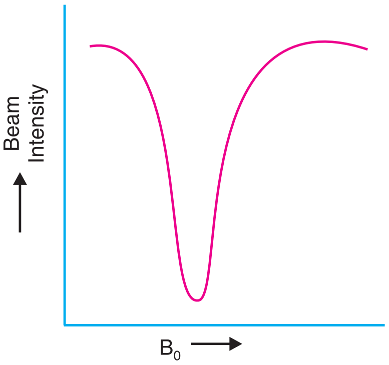


Fig. 19.23

In these experiments, the nuclear g_I factor is measured from the experimental data of the resonance frequency corresponding to the applied magnetic field.

$$g_I = \frac{h\upsilon_0}{\mu_N B}$$

It is found that $g_I = 2.167$.

The magnetic moment is related to nuclear g-factor by the following relation

$$\mu_I = g_I \mu_N I$$
$$= 2.167 \times (1 \text{ nuclear magneton}) \times 3/2 \qquad [I \text{ for Li}^7 = 3/2]$$
$$= 3.2505 \text{ nuclear magneton}$$

This method is particularly useful for magnetic moments of nuclei whose nuclear spins were accurately known.

19.17 Molecular Orbital Theory of Hydrogen Molecule Ion H_2^+

The hydrogen molecule ion H_2^+ consists of two hydrogen nuclei and one electron (Fig. 19.24). R is internuclear distance, $\boldsymbol{r_a}$ and $\boldsymbol{r_b}$ are distances of electron from nuclei $\boldsymbol{a}$ and b respectively.

The electron when in the neighbourhood of a is described by the atomic orbital Ψ_a centred on a. The electron when in the neighboarhood of b is described by the atomic orbital Ψ_b centred on b.

Ψ_a and Ψ_b are 1s atomic orbitals given by

$$\Psi_a = \frac{1}{\sqrt{\pi}}\left(\frac{1}{a_0}\right)^{3/2} e^{-r_a/a_0}, \qquad \Psi_b = \frac{1}{\sqrt{\pi}}\left(\frac{1}{a_0}\right)^{3/2} e^{-r_b \backslash a_0} \qquad \text{...(1)}$$

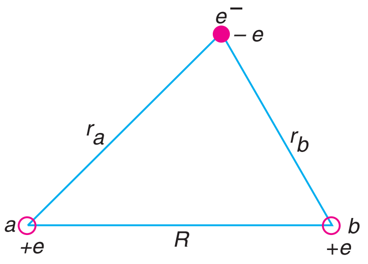


Fig. 19.24

a_0 is radius of Bohr's orbit.

A reasonable molecular orbital is a linear combination of Ψ_a and Ψ_b, i.e.,

$$\Psi = c_1\Psi_a + c_2\Psi_b \qquad \text{...(2)}$$

Here, c_1 and c_2 are constants.

This particular form of a linear combination of atomic 'orbitals' making a molecular orbital is called LCAO – MO for short.

The energy E of the system is given by the Schrodinger equation

$$H\Psi = E\Psi, \quad E = \frac{\langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle} \qquad \text{...(3)}$$

The Hamiltonian of the system is given by

$$H = \frac{-\hbar^2}{2m}\nabla^2 - \frac{e^2}{4\pi\varepsilon_0 r_a} - \frac{e^2}{4\pi\varepsilon_0 r_b} + \frac{e^2}{4\pi\varepsilon_0 \boldsymbol{R}} \qquad \text{...(4)}$$

Here, m is the mass of the electron.

Substituting the value of Ψ in Eq. (3), we get

$$E = \frac{\langle (c_1\Psi_a + c_2\Psi_b) | H | (c_1\Psi_a + c_2\Psi_b) \rangle}{\langle (c_1\Psi_a + c_2\Psi_b) | (c_1\Psi_a + c_2\Psi_b) \rangle}$$

Since the atomic orbitals Ψ_a and Ψ_b are normalized,

$$E = \frac{|c_1|^2 \langle \Psi_a | H | \Psi_a \rangle + c_1^* c_2 \langle \Psi_a | H | \Psi_b \rangle + c_2^* c_1 \langle \Psi_b | H | \Psi_a \rangle + |c_2|^2 \langle \Psi_b | H | \Psi_b \rangle}{|c_1|^2 + |c_2|^2 + c_1^* c_2 \langle \Psi_a | \Psi_b \rangle + c_2^* c_1 \langle \Psi_b | \Psi_a \rangle} \qquad \text{...(5)}$$

The integrals,

$$\langle \Psi_a | H | \Psi_a \rangle = H_{aa}, \qquad \langle \Psi_b | H | \Psi_b \rangle = H_{bb} \qquad \text{...(6a)}$$

are known as **Coulomb integrals.**

H_{aa} and H_{bb} are the energies of the atomic orbitals Ψ_a and Ψ_b in the molecular skeleton.

The integrals $\langle \Psi_a | H | \Psi_b \rangle = H_{ab}, \quad \langle \Psi_b | H | \Psi_a \rangle = H_{ba}$ ...(6b)

are known as **resonance** or **exchange integrals** and $H_{ab} = H^*_{ba}$.

The integral $\langle \Psi_a | \Psi_b \rangle = \langle \Psi_b | \Psi_a \rangle = S$ ...(6c)

is known as **overlap integral**.

Since the selection of the wave function Ψ is arbitrary, according to the variational theorem

$$E \geq E_0 \qquad \text{...(7)}$$

Here, E_0 is the exact ground state energy of the system.

Assuming the atomic orbitals Ψ_a and Ψ_b to be real, from Eq. (5),

$$E(c_1^2 + c_2^2 + 2c_1c_2S) = c_1^2 H_{aa} + c_2^2 H_{bb} + 2c_1c_2 H_{ab} \qquad ...(8)$$

In order to minimise E with respect to c_1 and c_2, it is necessary that

$$\frac{\partial E}{\partial c_1} = 0, \quad \text{and} \quad \frac{\partial E}{\partial c_2} = 0.$$

Applying the condition $\partial E/\partial c_1 = 0$ in Eq. (8), we get

$$(H_{aa} - E)c_1 + (H_{ab} - ES)c_2 = 0 \qquad ...(9)$$

The condition $\partial E/\partial c_2 = 0$ gives

$$(H_{ab} - ES)c_1 + (H_{bb} - ES)c_2 = 0 \qquad ...(10)$$

Equations (9) and (10) are known together as the **secular equations.**

These simultaneous homogeneous equations in c_1 and c_2 have non-trivial solutions only if

$$\begin{vmatrix} H_{aa} - E & H_{ab} - ES \\ H_{ab} - ES & H_{bb} - E \end{vmatrix} = 0 \qquad ...(11)$$

or $(H_{aa} - E)(H_{bb} - E) - (H_{ab} - ES)^2 = 0 \qquad ...(12)$

Since $H_{aa} = H_{bb}$, we get

$$H_{aa} - E = \pm (H_{ab} - ES) \qquad ...(13)$$

The two roots for E are the allowed energy values of the system.

$$E_1 = \frac{H_{aa} + H_{ab}}{1 + S}, \qquad E_2 = \frac{H_{aa} - H_{ab}}{1 - S} \qquad ...(14)$$

Combining Eqs. (9) and (13), we get

$$\pm (H_{ab} - ES)c_1 + (H_{ab} - ES)c_2 = 0 \qquad ...(15)$$

From Eq. (15), $c_1 = c_2$ for energy E_1 and $c_1 = -c_2$ for the energy E_2.

The wave functions, corresponding to energies E_1 and E_2 are respectively,

$$\Psi_1 = c_1 (\Psi_a + \Psi_b) \quad \text{and} \quad \Psi_2 = c_2 (\Psi_a - \Psi_b) \qquad ...(16)$$

Normalization gives

$$\Psi_1 = \frac{\Psi_a + \Psi_b}{\sqrt{2 + 2S}}, \qquad \Psi_2 = \frac{\Psi_a - \Psi_b}{\sqrt{2 - 2S}} \qquad ...(17)$$

The Coulomb integral

$$H_{aa} = \left\langle \Psi_a \left| \frac{-\hbar^2}{2m} \nabla^2 - \frac{e^2}{4\pi\varepsilon_0 r_a} \right| \Psi_a \right\rangle - \left\langle \Psi_a \left| \frac{e^2}{4\pi\varepsilon_0 r_b} \right| \Psi_a \right\rangle + \left\langle \Psi_a \left| \frac{e^2}{4\pi\varepsilon_0 R} \right| \Psi_a \right\rangle \qquad ...(18)$$

The first term on the right is simply the ground state energy of the hydrogen atom E_H since the operator in it is the hydrogen atom Hamiltonian and Ψ_a is the one electron wave function.

The nuclear repulsion term $\dfrac{e^2}{4\pi\varepsilon_0 R}$, is independent of the electronic coordinates.

Let

$$V_{aa} = \left\langle \Psi_a \left| \frac{e^2}{4\pi\varepsilon_0 r_b} \right| \Psi_a \right\rangle \quad \text{and} \quad V_{ab} = \left\langle \Psi_a \left| \frac{e^2}{4\pi\varepsilon_0 r_a} \right| \Psi_b \right\rangle \qquad ...(19)$$

$$H_{aa} = E_H - V_{aa} + \frac{e^2}{4\,\pi\varepsilon_0 R} \qquad ...(20)$$

$$H_{ab} = E_H S - V_{ab} + \frac{e^2}{4\,\pi\varepsilon_0 R} S \qquad ...(21)$$

The value of the quantities V_{aa}, V_{ab}, S and the nuclear repulsion energy $\frac{e^2}{4\pi\varepsilon_0 R}$ depend on the internuclear distance R and are always positive. The overlap $S = 0$, if the two nuclei are infinitely separated and $S = 1$ if they are together.

Substituting Eqs. (20) and (21) in Eq. (14), we get

$$E_1 = E_H - \frac{V_{aa} + V_{ab}}{1 + S} + \frac{e^2}{4\pi\varepsilon_0 R} \qquad \text{...(22)}$$

and

$$E_2 = E_H - \frac{V_{aa} - V_{ab}}{1 - S} + \frac{e^2}{4\pi\varepsilon_0 R} \qquad \text{...(23)}$$

Fig. 19.25 shows the relative energies of molecular orbitals and their constituent atomic orbitals. The original energy state of two hydrogen atoms possess the same energy E_H. But in H_2^+ ion they split into two new energy states, one of lower energy (symmetric molecular state) and the other of higher energy (anti symmetric molecular state) than the original 1 – S state. These two states are sometimes referred to as *bonding* and *antibonding* states of the hydrogen-molecule ion.

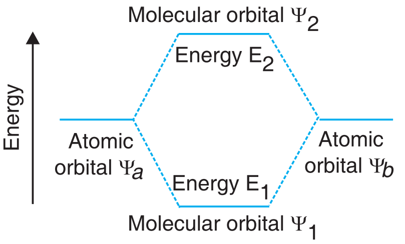

Molecular orbital Ψ2
Energy E2
Energy
Atomic orbital Ψa
Atomic orbital Ψb
Energy E1
Molecular orbital Ψ1

Fig. 19.25

Fig. 19.26 shows the probability density for the bonding and antibonding states.

Fig. 19.26(a) shows the probability density functions. Full lines represent Ψ_1^2 or Ψ_2^2. Broken lines represent Ψ_a^2 or Ψ_b^2.

Fig. 19.26(b) shows contours of equal electron density for Ψ_1^2 and Ψ_2^2.

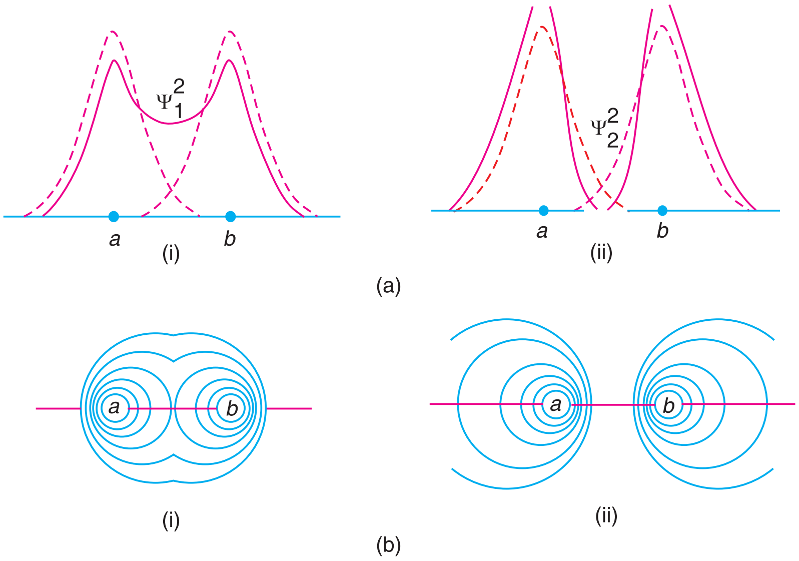

Ψ1²
a
b
(i)
Ψ2²
a
b
(ii)
(a)
a
b
(i)
a
b
(ii)
(b)

Fig. 19.26

Fig. 19.27 shows the binding (potential) energy of H_2^+ ion as a function of internuclear distance R. Antisymmetric orbital leads to an *unstable repulsive state.* Symmetric orbitals give the binding energy curve with a minimum which indicates the formation of a *stable state*. The minimum in the symmetric curve corresponds to the equilibrium distance of H^+ and H, $R = 2.5\ a_0$. This is the internuclear distance at which H_2^+ is most stable. The dissociation energy of H_2^+ ion is 1.77 eV. The experimental value for dissociation energy is 2.789 eV. A better agreement between theory and experiment is obtained by introducing a variational parameter in the trial wave function.

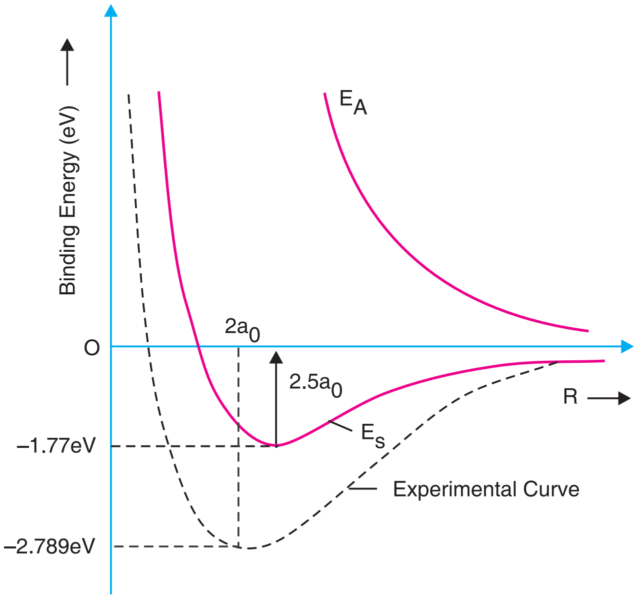


Fig. 19.27

19.18 Heitler-London Theory of Hydrogen Molecule

Fig. 19.28 shows the Hydrogen molecule system. It consists of two positively charged hydrogen nuclei a and b and two electrons whose coordinates are denoted by symbols 1, 2. r_{12} is the distance between two electrons 1 and 2. R is the internuclear distance. When **R** is large, the system may be regarded as consisting of two separated hydrogen atoms.

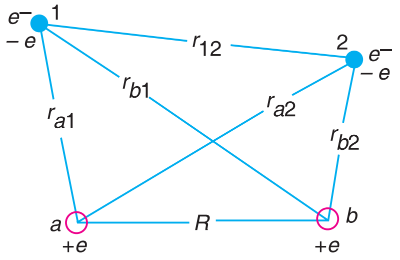


Fig. 19.28

The Schrodinger equations are

$$H_a(1)\ \Psi_a(1) = E_a\ \Psi_a(1) \quad H_b(2)\ \Psi_b(2)\ E_b\ \Psi_b\ (2) \quad ...(1)$$

Here, $E_a = E_b = E_H$, the ground state energy of the hydrogen atoms and $\Psi_a(1)$ and $\Psi_b(2)$ are the 1s hydrogenic wave functions.

$$H_a(1) = -\frac{\hbar^2}{2m}\nabla_1^2 - \frac{e^2}{4\pi\in_0 r_{a1}}, \quad H_b(2) = -\frac{\hbar^2}{2m}\nabla_2^2 - \frac{e^2}{4\pi\in_0 r_{b2}} \quad ...(2)$$

When the two atoms are brought closer and if there is no interaction between the two, the total Hamiltonian of the system.

$$H = H_a(1) + H_b(2) \quad ...(3)$$

It amounts to assuming that electron 1 is moving about proton a and electron 2 is moving about proton b.

$$\begin{aligned}[H_a(1) + H_b(2)]\ \Psi_a(1)\Psi_b(2) &= H_a(1)\Psi_a(1)\ \Psi_b(2) + H_b(2)\ \Psi_a(1)\ \Psi_b(2)\\ &= E_a\Psi_a(1)\ \Psi_b(2) + E_b\Psi_a(1)\ \Psi_b(2)\\ &= (E_a + E_b)\ \Psi_a(1)\ \Psi_b(2) \quad ...(4)\end{aligned}$$

So this system of two hydrogen atoms can be described by the wave function

$$\Psi_1(1,2) = \Psi_a(1)\Psi_b(2) \quad ...(5)$$

with energy eigenvalue $E_a + E_b$.

If the electron 1 is associated with atom b and electron 2 is associated with atom a, an equally good wave function with the same energy is

$$\Psi_2(2,1) = \Psi_a(2)\Psi_b(1) \qquad ...(6)$$

$\Psi_1(1,2)$ and $\Psi_2(2,1)$ are eigenfunctions of the Hamiltonian $(H_a + H_b)$ with eigenvalue $(E_a + E_b)$. The two functions differ only in the interchange of the two electrons between the obitals. Hence, this degeneracy is called the **exchange degeneracy.** The wave functions of the two electron system, must be a linear combination of $\Psi_1(1,2)$ and $\Psi_2(2,1)$.

The Hamiltonian operator for the hydrogen molecule is

$$H = \left(-\frac{\hbar^2}{2m}\nabla_1^2 - \frac{e^2}{4\pi\in_0 r_{a1}} - \frac{e^2}{4\pi\in_0 r_{b1}}\right) + \left(-\frac{\hbar^2}{2m}\nabla_2^2 - \frac{e^2}{4\pi\in_0 r_{b2}} - \frac{e^2}{4\pi\varepsilon_0 r_{a2}}\right)$$

$$+ \frac{e^2}{4\pi\in_0 r_{12}} + \frac{e^2}{4\pi\varepsilon_0 R} \qquad ...(7)$$

or

$$H = H_a(1) + H_b(2) + H' + \frac{e^2}{4\pi\varepsilon_0 R} \qquad ...(8)$$

Here,

$$H' = -\frac{e^2}{4\pi\varepsilon_0 r_{b1}} - \frac{e^2}{4\pi\varepsilon_0 r_{a2}} + \frac{e^2}{4\pi\varepsilon r_{12}} \qquad ...(9)$$

An exchange of the two electrons leaves the Hamiltonian H unchanged. So the wave functions must either be symmetric or antisymmetric with respect to such an exchange.

The symmetric ψ_s and antisymmetric ψ_{as} combinations are

$$\Psi_s = N_s\,[\Psi_a(1)\,\Psi_b(2) + \Psi_a(2)\,\Psi_b(1)] \qquad ...(10)$$

$$\Psi_{as} = N_a\,[\Psi_a(1)\,\Psi_b(2) - \Psi_a(2)\,\Psi_b(1)] \qquad ...(11)$$

Here, N_s and N_a are normalization constants.

Normalization condition gives

$$N_s^2 = \frac{1}{2(1+S^2)}, \qquad N_a^2 = \frac{1}{2\,(1-S^2)} \qquad ...(12)$$

Here,

$$S = \langle \Psi_a \mid \Psi_b \rangle = \langle \Psi_b \mid \Psi_a \rangle \qquad ...(13)$$

The inclusion of electron spin and Pauli's principle leads to the Heitler–London wave functions

$$N_s[\Psi_a(1)\,\Psi_b(2) + \Psi_a(2)\,\Psi_b(1)]\,\frac{1}{\sqrt{2}}\,[\alpha(1)\,\beta(2) - \beta(1)\,\alpha(2)] \qquad ...(14)$$

$$N_a[\Psi_a(1)\,\Psi_b(2) - \Psi_a(2)\,\Psi_b(1)]\begin{cases}\alpha(1)\,\alpha(2)\\ \dfrac{1}{\sqrt{2}}\,[\alpha(1)\,\beta(2) + \alpha(2)\,\beta(1)]\\ \beta(1)\,\beta(2)\end{cases} \qquad ...(15)$$

The first one corresponds to a singlet ($S = 0$) and the second one to a triplet (S = 1). As the Hamiltonian does not contain spin terms, the energy is not affected by the inclusion of spin part. The space parts alone can then be taken as the unperturbed wave functions for the evaluation of energy.

A perturbation calculation is then carried out with H' (Eq. 9) as perturbation. The first order correction to the energy E' is the diagonal matrix element of the perturbing Hamiltonian corresponding to the unperturbed wave functions. Then

$$E_1' = N_S^2 \langle \Psi_a(1)\Psi_b(2) + \Psi_a(2)\Psi_b(1) \,|\, H' \,|\, (\Psi_a(1)\Psi_b(2) + \Psi_a(2)\Psi_b(1)) \rangle$$

$$= \frac{(2J+2K)}{2(1+S^2)} = \frac{J+K}{1+S^2} \qquad \text{...(16)}$$

Here, $$J = \langle \Psi_a(1)\Psi_b(2) \,|\, H' \,|\, \Psi_a(1)\,\Psi_b(2) \rangle \qquad \text{...(17)}$$

J is called *coulomb integral*.

$$K = \langle \Psi_a(1)\Psi_b(2) \,|\, H' \,|\, \Psi_a(2)\,\Psi_b(1) \rangle \qquad \text{...(18)}$$

K is called *exchange integral*.

The energy of the singlet state corrected to first order

$$E_S = 2E_H + \frac{J+K}{1+S^2} + \frac{e^2}{4\pi\varepsilon_0 R} \qquad \text{...(19)}$$

Singlet state is the *stable state*.

The first order correction to the triplet state

$$E_2' = \frac{J-K}{1-S^2} \qquad \text{...(20)}$$

The energy of the triplet state

$$E_a = 2E_H + \frac{J-K}{1-S^2} + \frac{e^2}{4\pi\varepsilon_0 R} \qquad \text{...(21)}$$

Experimental and theoretical results :

Fig. 19.29 shows the plots of energies E_1' and E_2' against internuclear distance R. The equilibrium internuclear distance r_0 is the one corresponding to the minimum of the Ψ_s curve. The theoretical value of 0.085 nm for r_0 is high compared to the experimental value of 0.074 nm.

The state characterised by the Ψ_{as} combination corresponds to repulsion for all values of R.

Fig. 19.30. shows distribution of charge density of electron cloud in (*a*) stable and (*b*) unstable states of a hydrogen molecule.

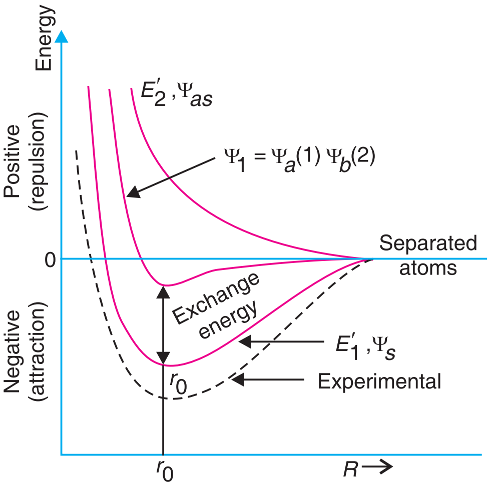


Fig. 19.29

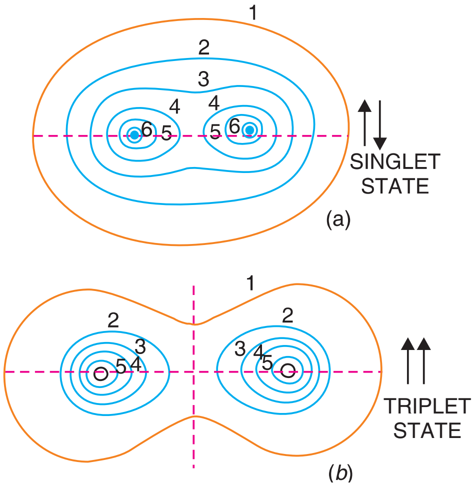
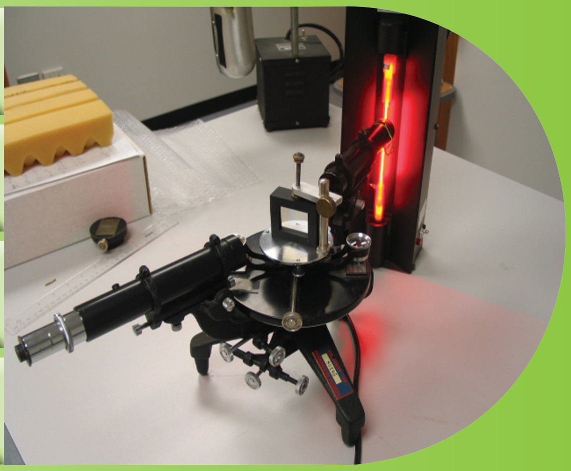


Fig. 19.30

CHAPTER
20



Fine Structure of H_α Line

20.1 Fine Structure of H_α Line

According to Bohr's theory, H_α line, the first of Balmer series, arises due to transition from the third quantum state ($n = 3$) to the second quantum state ($n = 2$).

Fig. 20.1 shows the different energy states for $n = 2$ and $n = 3$ in case of Hydrogen atom.

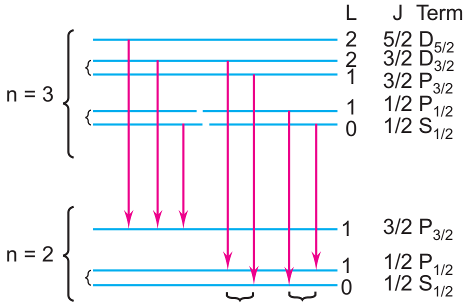

L
J Term
n = 3
2 5/2 D5/2
2 3/2 D3/2
1 3/2 P3/2
1 1/2 P1/2
0 1/2 S1/2
n = 2
1 3/2 P3/2
1 1/2 P1/2
0 1/2 S1/2

Fig. 20.1

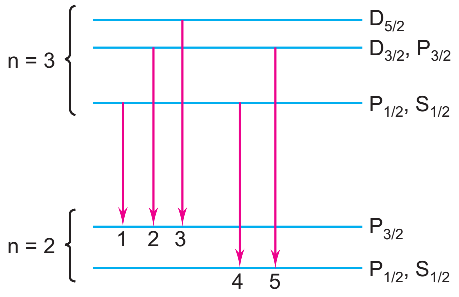

n = 3
D5/2
D3/2, P3/2
P1/2, S1/2
n = 2
P3/2
P1/2, S1/2
1
2
3
4
5

Fig. 20.2

For the upper state $n = 3$

L values	0	1	2
$J = L \pm S$	$\frac{1}{2}$	$\frac{3}{2}$ and $\frac{1}{2}$	$\frac{5}{2}$ and $\frac{3}{2}$ $\left(\because S = \frac{1}{2}\right)$
Spectral terms	$^2S_{1/2}$	$^2P_{3/2}$ and $^2P_{1/2}$	$^2D_{5/2}$ and $^2D_{3/2}$

For the lower state $n = 2$

L values	0	1
$J = L \pm S$	$\frac{1}{2}$	$\frac{3}{2}$ and $\frac{1}{2}$
Spectral terms	$^2S_{1/2}$	$^2P_{3/2}$ and $^2P_{1/2}$

Theoretically fifteen transitions are possible between the five terms of the upper state ($n = 3$) and three terms of the lower state ($n = 2$). But by applying the selection rules namely, $\Delta L = \pm 1$ and $\Delta J = 0$ or ± 1, these transitions reduce to the following seven transitions:

$$3\,{}^2D_{5/2} \to 2\,{}^2P_{3/2},\quad 3\,{}^2D_{3/2} \to 2\,{}^2P_{3/2},\quad 3\,{}^2D_{3/2} \to 2\,{}^2P_{1/2},\quad 3\,{}^2P_{3/2} \to 2\,{}^2S_{1/2},$$
$$3\,{}^2P_{1/2} \to 2\,{}^2S_{1/2},\quad 3\,{}^2S_{1/2} \to 2\,{}^2P_{3/2},\quad 3\,{}^2S_{1/2} \to 2\,{}^2P_{1/2}.$$

Out of these seven allowed transitions, two pairs in separate cases are identical, *i.e.,*

$$[3\,{}^2D_{3/2} \to 2\,{}^2P_{1/2} \quad \text{and} \quad 3\,{}^2P_{3/2} \to 2\,{}^2S_{1/2}];$$
$$[3\,{}^2P_{1/2} \to 2\,{}^2S_{1/2} \quad \text{and} \quad 3\,{}^2S_{1/2} \to 2\,{}^2P_{1/2}]$$

since they represent transitions between coincident levels *i.e.,* levels whose L values differ by unity, but J values are the same.

Thus the fine structure of H_α line should have five components.

Now applying the intensity rules, the transitions $D_{5/2} \to P_{3/2}$ and $[D_{3/2} \to P_{1/2}$ and $P_{3/2} \to S_{1/2}]$ have the maximum intensity because they are of type $\Delta L = -1$ and $\Delta J = -1$. The two components $D_{3/2} \to P_{3/2}$ and $P_{1/2} \to S_{1/2}$ are less intense since they are of the type $\Delta L = -1$ and $\Delta J = 0$. The components $S_{1/2} \to P_{3/2}$ is still weaker since here $\Delta L = +1$ and $\Delta J = +1$. The weakest transition is $S_{1/2} \to P_{1/2}$ because here $\Delta L = +1$ and $\Delta J = 0$.

Energy level diagram. The various energy levels are drawn in as the horizontal levels as shown in Fig. 20.1. The coincident levels are drawn close to each other. There are two such levels in upper state and one in the lower state. All the seven allowed transitions are also shown in Fig. 20.1. Leaving the transitions between coincident levels, the five transitions are separately shown in Fig. 20.2.

CHAPTER
21

FluoroMax-2

ELECTRONIC SPECTRA OF MOLECULES

AT A GLANCE

21.1	Electronic Spectra of Molecules
21.2	Fluorescene and Phosphorescence

21.1 Electronic Spectra of Molecules

The electronic spectra of molecules arise when the electrons of a molecule are excited to higher energy states. The energy involved in this process is generally large. Therefore, the electronic spectra of molecules are usually found in visible and ultraviolet region. Thus electronic spectra arise due to the changes in the arrangement of molecular electrons. The molecule is characterised by a discrete set of electronic energy states.

We can write the total energy E of a molecule as the sum

$$E = E_e + E_v + E_j, \qquad \text{...(1)}$$

where E_e is the energy associated with the electronic exciation, and E_v and E_j are the vibrational and rotational energies. In a transition, the total energy change is the sum

$$E_i - E_f = (E_{e_i} - E_{e_f}) + (E_{v_i} - E_{v_f}) + (E_{j_i} - E_{j_f})$$

or
$$\Delta E = \Delta E_e + \Delta E_v + \Delta E_j,$$

where ΔE_e, ΔE_v and ΔE_j represent the changes in the electronic, vibrational, and rotational energies respectively. These are of the order of

$\Delta E_e \sim 1$ eV, $\Delta E_v \sim 0.1$ eV $\Delta E_j \sim 0.001$ eV.

For a given electronic transition, the spectrum consists of many lines arising due to the energy changes ΔE_v and ΔE_j. These lines are so close to each other in frequency that the spectrum looks like a band.

Consider the vibrational structure of a band due to a transition between electronic states F_{e_2} and F_{e_1}. The probability of transition between two given vibrational levels of two electronic states is determined by *Franck Condon principle*.

According to Franck-Condon principle, the transition between different vibrational states should be represented by vertical lines. This is due to the fact that electrons move and rearrange themselves much faster than the nuclei of molecules. For example, the time for an electron to circle a hydrogen nucleus is of the order of 10^{-16} s, which is thousand times less than the typical period of vibration 10^{-13} s of the molecule. Thus the internuclear distance (R) of a vibrating molecule does not change appreciably in course of transition between two electronic states. *Hence the transition should be represented by a vertical line*, that is a line of constant R.

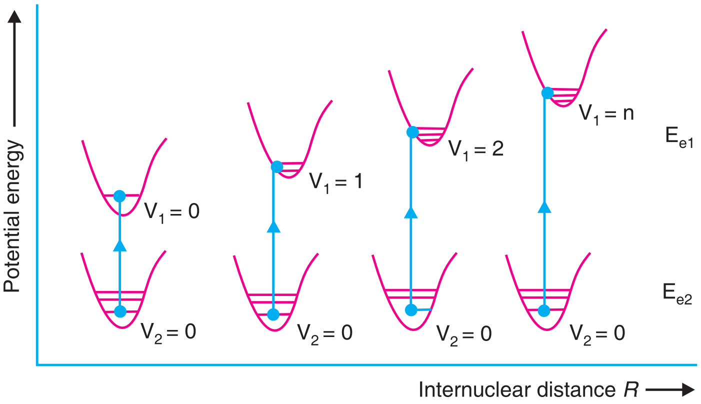


Fig. 21.1

Secondly, the Franck-Condon principle states that transitions from one state to another are most probable when nuclei are in their mean positions *i.e.*, *observed transitions between two states should start from extreme positions of vibrational level.* This is due to the fact that nuclei spend longest time at mean positions as the vibrational kinetic energy is zero there. Fig. 21.1 shows the potential energy of the molecules as a function of internuclear distance for two electronic energies F_{e_1} and F_{e_2} with vibrational energy levels. Quantum mechanics predicts that most probable position of the nuclei for the lowest vibrational energy ($v_2 = 0$) is the equilibrium position *i.e.*, mid-position. However, the most probable internuclear distance for vibrational levels other than $v_2 = 0$, corresponds to the extreme positions. Thus the transitions start from extreme positions for levels other than $v_2 = 0$.

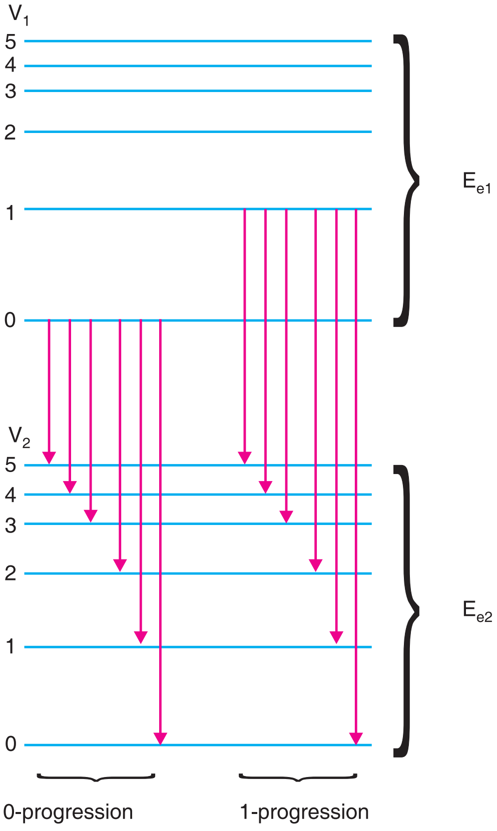


Fig. 21.2

As the equilibrium values of internuclear distances in a molecule will be different in different electronic states, the vibrational frequencies will be different in different states. Again considering the fact that the vibrations of the molecule are not simple harmonic, *i.e.*, vibrations are anharmonic vibrations, the lines get crowded towards the high frequency end of the spectrum. The electronic emission spectrum of a molecule due to vibration is shown in Fig. 21.2. Each of these lines is spread out into a band due to transitions between different rotational energy levels. The bands in a band system are grouped together into different progressions. Each progression consists of the set of bands which arise from transitions from a fixed initial vibrational energy level to different final vibrational levels. In Fig. 21.2, zero and one progressions are shown. If the progression obtained by transitions from vibrational energy in upper electronic state to different vibrational energy levels in lower electronic state are considered, we find that the progression ends in a continuum. This continuum arises due to dissociation of the molecule.

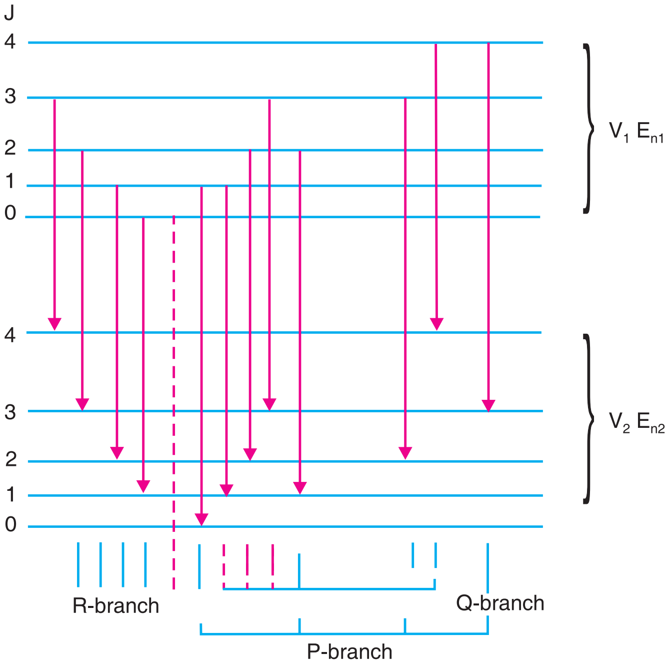


Fig. 21.3

Now we consider the rotational lines in each band. These lines are again not equally spaced due to the fact that moments of inertia of molecules are different in different electronic states. Thus the lines in a band get crowded towards one end. The intensity also shows an abrupt change. There are no selection rules governing the changes between vibrational energy levels. The transitions between rotational energy levels obey the selection rule $\Delta J = -1, 0$ or $+1$ with additional requirement that transition with $J = 0$ to another level with $J = 0$ is not allowed. In this case, we get three branches of lines in any given band namely the *R*-branch, *Q*-branch and *P*-branch (Fig. 21.3).

21.2 Fluorescene and Phosphorescence

Molecules which absorb radiation in the near ultraviolet and visible range may re-emit it at a longer wavelength. This phenomenon is called *fluorescence* and is illustrated in Fig. 21.4.

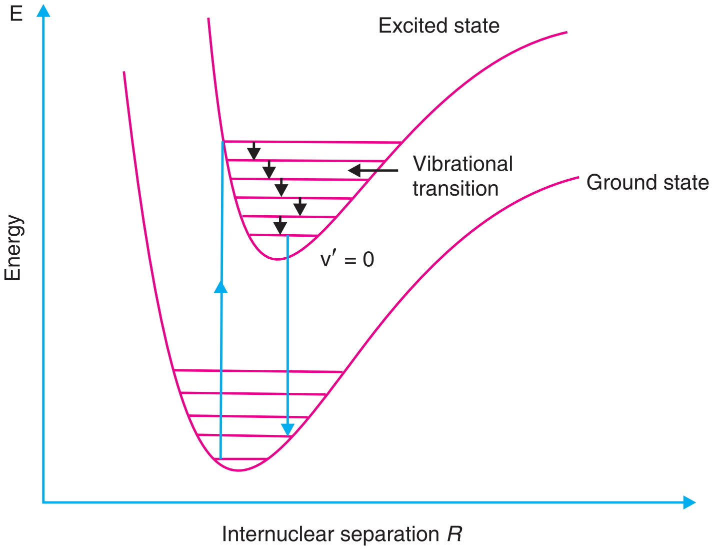


Fig. 21.4

First, absorption transitions occur from a lower electronic state to some vibrational levels of an upper electronic state. The absorption transitions are governed by the Franck-Condon principle. For example, we have shown in Fig. 21.4 a transition leading from the lowest vibrational level of the ground electronic state of the molecule to an excited vibrational level of an upper electronic state. The molecule may then lose vibrational energy through collisions with other molecules (thermal decay) and reach the lowest vibrational state $v' = 0$ of the excited electronic level. This process is often more rapid than spontaneous emission of radiation, so that when the molecule finally re-emits radiation, it is from the vibrational level $v' = 0$. Fluorescent radiation is, therefore, of lower frequency than that of the absorbed radiation.

Fluorescence Imaging System.

Fluorescence excited by ultraviolet light has many applications, for instance to help identify minerals and biochemical compounds. This principle is applied in fluorescent lamp.

The fluorescence phenomenon involves a spontaneous emission between two electronic states of the *same multiplicity*. But *phosphorescence* phenomenon involves the spontaneous emission from an excited electronic state to a lower one of *different multiplicity*. For example, let us assume that a molecule first undergoes an absorption transition from a singlet ground state ^1X to a singlet excited state ^1A (Fig. 21.5).

Thermal degradation of the vibrational energy then occurs via collisions with other molecules. Suppose now that the state ^1A is coupled to a triplet state ^3A of similar energy. Instead of going down the vibrational ladder of the state ^1A, the molecule may then undergo an intersystem crossing and be transferred by a radiationless transition into the triplet state ^3A. After the crossing has occurred, thermal decay will proceed within the ^3A well, until the molecule has reached the lowest vibrational level of the ^3A state. This triplet state may now decay to the ground (singlet) state ^1X by a radiative transition called phosphorescent emission. Indeed, although electric dipole transitions between the excited ^3A state and the ground state ^1X are spin-forbidden because of the selection rule $\Delta S = 0$, radiative transitions $^3A \rightarrow {}^1X$ may not be absolutely forbidden, and may occur slowly through spin -orbit interactions. Such transitions accordingly have long half-lives, and the resulting phosphorescent radiation may be emitted minutes or even hours after the initial absorption.

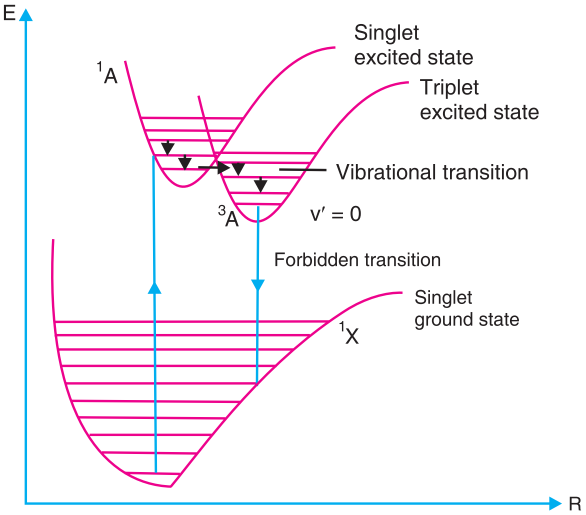
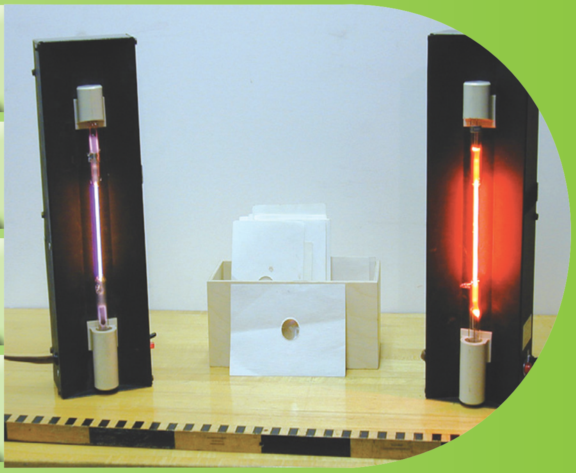


Fig. 21.5

Phosphorescence

EXERCISE

1. Give an elementary account of the origin of electronic spectra of molecules.
2. Write a short note on Franck-Condon principle. **(P.U. 93)**



CHAPTER 22

ATOMIC SPECTRA

AT A GLANCE

22.1	Finite Width of Spectral Line
22.2	The Widths of Spectral Lines

22.1 Finite Width of Spectral Line

Consider the radiation of light from an 'excited' atom (Fig. 22.1). The average life time of an excited state (energy level) is about 1×10^{-8} second. At the most, the uncertainty in measurement of life time is of the order of their life time, *i.e.*, $\Delta t = 1 \times 10^{-8}$ s. According to the uncertainty principle, there will be a corresponding energy uncertainty (width) ΔE of this energy level.

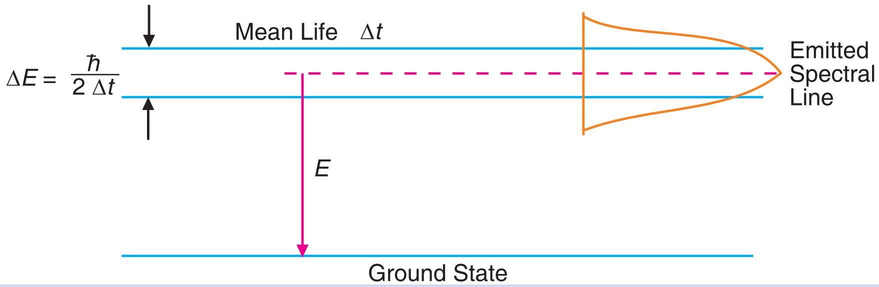


Fig. 22.1

$$\Delta E = \frac{\hbar}{2\,\Delta t} = \frac{1\cdot 05\times 10^{-34}}{2\times 10^{-8}}\,\mathbf{J} = 5\cdot 25\times 10^{-27}\,\mathbf{J}.$$

This means, that excited energy levels have a finite energy spread so that the radiation emitted when atoms fall back to the ground state from a particular level is not truely monochromatic. In other words, the spectral lines can never be infinitely sharp but must have a natural finite width (natural broadening).

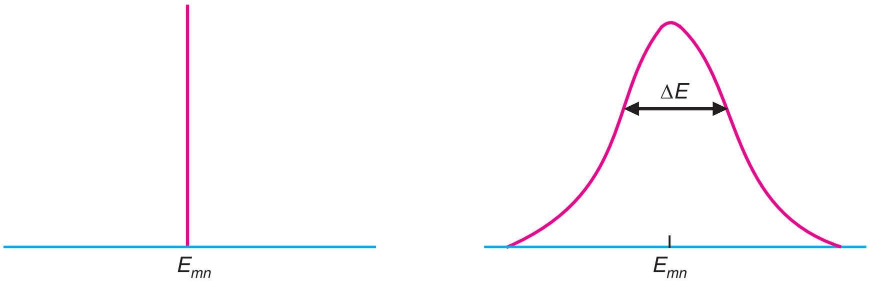
The frequency of light is uncertain by the amount

$$\Delta\nu = \frac{\Delta E}{h} = \frac{5 \cdot 25 \times 10^{-27}}{6 \cdot 625 \times 10^{-34}} = 7 \cdot 925 \times 10^{6} \text{ Hz.}$$

Thus, the spectral line cannot be a sharp geometrical line but must be a line of width given by ($\Delta\nu$) hertz. Thus, the natural line width is due to the uncertainty principle.

22.2 The Widths of Spectral Lines

In the spectroscopic measurements, the instruments used influence the sharpness of spectrum lines because of their limited resolving power. In the observations for spectral lines with the aid of low resolving power instruments, this is the principal cause for the broadening of lines. But up-till now, it has been established beyond doubt that no spectrum line is strictly sharp, but has a finite width, no matter how great the resolving power of the instrument is (Fig. 22.2). A line devoid of structure appears densest in the centre and fades out on the edges. For convenience, *the width of a line is defined as "the distance between two points, one on each side of the centre, at which the intensity is half as great as it is in the centre of this line."* This is sometimes called the half intensity width. The known causes of line broadening are :



(*a*) (*b*)

Fig. 22.2

(*a*) Spectral line having no width.

(*b*) Spectral line having width ΔE at half height.

1. The natural width, 2. Doppler broadening, 3. Pressure broadening.

(*i*) Doppler broadening

Consider an electrical discharge tube containing gas, acting as the source of light. Due to thermal agitation, most of the atoms emitting light have high velocities. Doppler effect modifies the apparent frequency of radiation from atoms moving with random thermal velocities.

The natural frequency of the spectral line is due to the atoms having no velocity component in the direction of the observer.

Doppler broadening is easily calculated for a gas where the only motions are the random motions of the atom associated with the thermal energy of the gas. The translational motion of an atom emitting radiation produces a change in the frequency given by expression

$$\frac{\Delta\nu}{\nu_0} = \frac{\nu - \nu_0}{\nu_0} = \frac{v \cos\theta}{c} = \frac{u}{c} \qquad \text{...(1)}$$

Here, v is the velocity of the atom,

θ is the angle between v and direction of observation,

ν_0 is the frequency of the line for $v = 0$,

ν is the observed frequency and

$u = v \cos \theta$ = the component of v in the direction of observation.

If the particle is approaching the observer with the component velocity u, the frequency becomes $\nu_0\left(1+\frac{u}{c}\right)$, and if it is going away, the frequency becomes $\nu_0\left(1-\frac{u}{c}\right)$.

In terms of wavelength, if a source of light of wavelength λ is moving in the line of sight with velocity u relative to the observer, the apparent change in wavelength as measured by the observer is

$$d\lambda = \frac{\lambda u}{c} \qquad \text{...(2)}$$

The velocity of motion (u) of the radiating atom can be found from the kelvin temperature T of the source. If m is the mass of the atom and k the Boltzmann constant,

$$\frac{1}{2}mu^2 = \frac{3}{2}kT$$

$$\therefore \qquad u = \sqrt{\frac{3kT}{m}}$$

Eq. (2) becomes,

$$d\lambda = \frac{\lambda}{c}\cdot\sqrt{\frac{3kT}{m}} \qquad \text{...(3)}$$

Eq. (3) shows that Doppler broadening is

(*i*) proportional to the square root of absolute temperature,

(*ii*) inversely proportional to the square root of the atomic weight.

To produce sharper lines in any given spectrum, the temperature of the source must be lowered. For example, low temperature source such as a liquid-air-cooled gaseous discharge tube can be used to get sharp lines.

The Doppler broadening is least with atoms of high atomic weight and this suggests that Hg^{198} *should be an ideal light source for sharp spectral lines.*

Thus, the Doppler broadening is due to random motion of the atoms in the source of light.

(*ii*) Pressure broadening

The life times of energy levels may be influenced by external conditions such as pressure. At the time of emission or absorption the atoms are under heavy pressure and the rate of collisions increases. If the atom at the time of emission of a radiation undergoes a collision process, the life time of the excited state will change. According to the uncertainty principle, the energy spread (ΔE) of a level will be increased. Thus, the effect of collision is to broaden the width of a spectral line. The width caused by this process may greatly exceed the natural line width. This process is referred to as pressure broadening. The pressure broadening of the lines depends on the density, temperature and nature of the gas within which the atom is situated.

In order to reduce the line width increase produced by collisions, the light source used in optical spectra research is often in a state of low pressure (such as 1 mm of Hg).

CHAPTER
23

MOLECULAR SPECTRA

AT A GLANCE

23.1 Origin of Molecular Spectra
23.2 Nature of Molecular Spectra
23.3 Different Modes of Molecular Excitation
23.4 Factors Affecting Line Width of Molecular Spectra
23.5 Factors Affecting Intensity of Molecular Spectra
23.6 Born-Oppenheimer Approximation
23.7 Rotation of Linear System
23.8 Rotational Spectra of Rigid Linear Molecules (See Section 19.7)
23.9 Non-Rigid Rotator
23.10 The Energy of a Diatomic Molecule
23.11 Vibrating Diatomic Molecule as a Harmonic Oscillator
23.12 Infra-red Spectra
23.13 Theory of the Origin of the Vibration-Rotation Spectrum of a Molecule
23.14 The Anharmonic Oscillator
23.15 Fundamental and Overtone Frequencies
23.16 Calculation of Force Constant

23.1 Origin of Molecular Spectra

The molecular spectra arise due to transitions between allowed energy states of molecules.

Molecular spectra (band spectra) are produced due to excitation of a substance in the *molecular state.*

The transition between two energy states of a molecule, as a result of absorption or emission of energy will give rise to a spectral line. When the molecule suffers a transition from higher energy state with energy E to a lower energy state with energy E', the difference of energy $\Delta E = (E - E')$ is emitted as radiation of frequency ν given as

$$\nu = \frac{E - E'}{h}.$$

For each allowed transition, we get one line in the spectrum of the molecule. Similarly we have the frequency in the absorption spectrum when the molecule suffers a transition from lower energy state to higher energy state.

AUSIM
DANGER
DO NOT PUT FINGERS IN SOCKETS WHEN TURNED ON

Molecular Spectra of two Element Molecules.

23.2 Nature of Molecular Spectra

Molecular spectrum is a band spectrum. Usually a band has a sharp intense edge on one side called '*band head*' and gradually decreases in intensity towards the other side.

With high resolving power instruments, band spectra are found to have threefold structure:

(*i*) Each band is composed of a large number of lines, which are crowded together at the long wavelength side called 'band head'. Going away from the band head, the lines become more widely spaced and also weaker.

(*ii*) There is a regular sequence in which several bands follow one another, and thus constitute a group of bands.

(*iii*) These different groups of bands are quite close and thus because of their regular arrangement, form a band system.

23.3 Different Modes of Molecular Excitation

When a molecule is supplied with energy, it may be excited in any of the following modes :

1. The molecule as a whole may rotate about an axis passing through the centre of mass (Fig. 23.1).

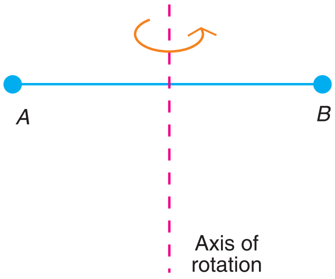


Fig. 23.1

2.The molecule may make vibrational motion. *i.e.*, the constituent atoms may vibrate relative to one another (Fig. 23.2). In a diatomic molecule, the two atoms usually vibrate along the line joining them, somewhat like a linear harmonic oscillator. We may think of a molecule as an arrangement of masses held together by massless springs. The masses represent the atomic nuclei and springs represent interatomic forces.

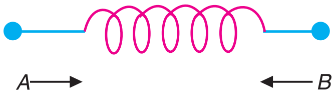


Fig. 23.2

3.There may be changes produced in its electronic configuration.

23.4 Factors Affecting Line Width of Molecular Spectra

When a molecule makes transition from a higher energy state to a lower energy state, molecular emission spectrum is produced. The spectrum is found to have some line-width, even if all attempts are made to narrow down the slit.

The factors that affect the line width of molecular spectra are the natural line width (due to the uncertainty principle), Doppler effect at moderate frequencies and intermolecular collisions (collision broadening) at lower frequencies (Sections 22.2 and 22.3).

Ideally, the analysis of any spectrum requires that the lines be well separated or resolved. In addition to the various mechanisms that can cause line broadening, lines can also be “broadened” by the measuring instrument itself. We define *resolving power* of an instrument as a measure of its ability to distinguish lines that overlap. It is usually expressed as the ratio of the observed wavelength (frequency) to the smallest difference between two wavelengths (frequencies) that can be measured, that is, $\lambda / \Delta\lambda$ ($\nu / \Delta\nu$) (Fig. 23.3).

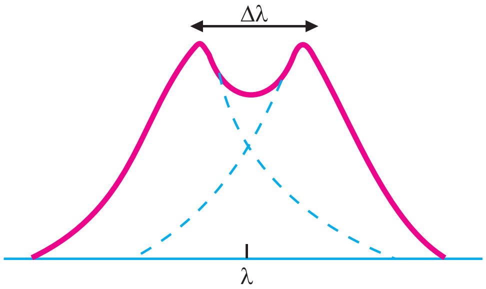


Fig. 23.3

Finally, rate processes such as dissociation, rotation, and electron and proton transfer reactions can also cause the broadening.

23.5 Factors Affecting Intensity of Molecular Spectra

The following three factors affect the intensity of molecular spectra.

1. *The transition probability. i.e.*, the likelihood of a system in one state changing to another state.

2. *The population of molecules. i.e.*, the number of molecules initially in the state from which the transition occurs. In general, the intensity of the absorption lines is proportional to the population in the lower state.

3. *The concentration or path length of the sample. i.e.*, the amount of material present, giving rise to the spectrum.

23.6 Born-Oppenheimer Approximation

So far we have considered rotation and vibration quite separately. The energies of the two motions are very different. We may, as a first approximation, consider that a diatomic molecule can execute rotations and vibrations quite independently. This is called the Born-Oppenheimer Approximation.

23.7 Rotation of Linear System

The simplest molecule is a diatomic molecule. *Example* : HCl, OCS. A diatomic molecule is an example of a linear system. In a diatomic molecule two atoms are bound by attractive forces. It can be regarded as a rigid dumb-bell. To a first approximation, it is assumed that the distance between two nuclei remains constant, *i.e.*, the molecule behaves as a *rigid rotator* (Fig. 23.4).

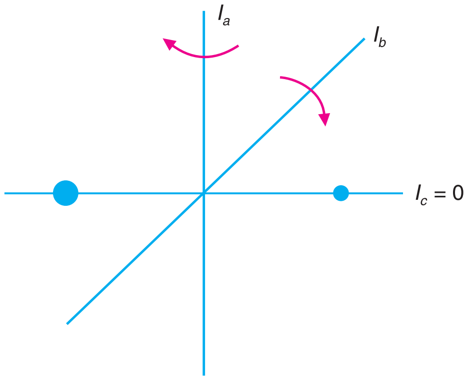


Fig. 23.4

The three directions of rotation may be taken as

(*i*) about the bond axis

(*ii*) end-over-end rotation in the plane of the paper and

(*iii*) end-over-end rotation at right angles to the plane.

For linear molecules, the moments of inertia of (*ii*) and (*iii*) are equal and that of (*i*) is zero.

$$I_a = I_b \text{ and } I_c = 0.$$

The principal axes of inertia are conventionally given the subscripts of *a*, *b*, and *c*, which may be thought of as the *x*, *y*, and *z* directions.

23.8 Rotational Spectra of Rigid Linear Molecules

(See Section 19.7)

Example 1. *The lines in the pure rotational spectrum of HCl are spaced as* 20.8×10^2 *per metre. Calculate the moment of inertia and the internuclear distance. Mass of proton* $= 1.67 \times 10^{-27}$ *kg ; mass of chlorine* $= 58.5 \times 10^{-27}$ *kg.*

Sol. The reduced *m'* is defined as

$$m' = \frac{m_1 m_2}{m_1 + m_2} = \frac{(1 \cdot 67 \times 10^{-27})(58 \cdot 5 \times 10^{-27})}{(1 \cdot 67 + 58 \cdot 5) \times 10^{-27}} = 1 \cdot 62 \times 10^{-27} \text{ kg}$$

Here, $$\Delta \bar{\nu} = \frac{\hbar}{2\pi Ic} = 20 \cdot 8 \times 10^2 \; m^{-1}$$

$$\therefore \quad I = \frac{\hbar}{2\pi c \Delta \bar{\nu}} = \frac{1 \cdot 054 \times 10^{-34}}{2\pi (3 \times 10^8)(20 \cdot 8 \times 10^2)} = 2 \cdot 689 \times 10^{-47} \text{ kg m}^2$$

$$\therefore \quad r = \sqrt{\frac{I}{m'}} = \sqrt{\frac{2 \cdot 689 \times 10^{-47}}{1 \cdot 62 \times 10^{-27}}} \text{ m} = 1 \cdot 286 \times 10^{-10} \text{ m.}$$

23.9 Non-Rigid Rotator

On the wave number scale, the frequency difference between two successive lines in the pure rotational spectrum of a diatomic molecule is given by

$$\Delta\bar{\nu} = \frac{\hbar}{2\pi Ic}$$

The rotational spectrum can be recorded. The absorption lines are equi-spaced. The separation between the adjacent lines is identified as $2B$.

$$\Delta\bar{\nu} = 2B = \frac{\hbar}{2\pi Ic}$$

$$\therefore \qquad B = \frac{\hbar}{4\pi Ic}\cdot\cdot$$

By measuring $\Delta\bar{\nu}$, the rotational constant B can be calculated. From this, the moment of inertia of the molecule I can be calculated.

$I = m' r^2$ where m' is the reduced mass of the molecule and r the bond length.

Knowing the value of m', we can calculate r.

Experimental investigations have shown that the successive lines in the far infra-red spectrum are not evenly spaced, but that the frequency separation decreases slightly with increasing value of J (larger speed of rotation). It shows that bond length (r) increases with J. Therefore, our assumption that the molecule is a rigid rotator is false. In fact, all bonds are elastic to some extent. The more quickly a diatomic molecule rotates, the greater is the centrifugal force tending to move the atoms apart.

Now we discuss the consequences of the change in bond length with J.

(*i*) When a bond is elastic, it will stretch and compress periodically with a certain functional frequency dependent upon the masses of the atoms and the elasticity (or force constant k) of the bond. This means that the molecule may have vibrational energy.

If the vibrational motion is simple harmonic, the force constant k is given by

$$k = 4\pi^2 \omega^2 c^2 m, \qquad \text{...(1)}$$

Here, ω is the vibration frequency (in m^{-1}). m is the reduced mass of the molecule.

The variation of B with J is determined by the force constant, *i.e.*, the weaker the bond, more readily will it distort under centrifugal forces.

(*ii*) The second consequence of elasticity is that the quantities r and B *vary during a vibration*.

Energy Levels

Consider a single particle of mass m rotating about a fixed point with an angular velocity, ω. Let the particle be at a distance r_0 from the fixed point when there is no rotation. Let this length increase to r, when the particle rotates.

Centrifugal force during rotation = $m r \omega^2$.

Restoring force due to bond stretching = $k(r - r_0)$.

The above two forces balance each other at any instant of rotation.

$$\therefore \qquad k(r - r_0) = m r \omega^2 \qquad \text{...(2)}$$

$$\text{or} \qquad r = \frac{k r_0}{(k - m\omega^2)} \qquad \text{...(3)}$$

This gives the distorted bond length.

Total energy of the rotating system = $K.E. + P.E.$

$$E = \frac{1}{2}I\omega^2 + \frac{1}{2}k(r-r_0)^2 = \frac{1}{2}I\omega^2 + \frac{1}{2}k\cdot\frac{m^2r^2\omega^4}{k^2}$$

$$\left[\because (r-r_0) = \frac{mr\omega^2}{k} \text{ from Eq. (2)}\right]$$

$$= \frac{1}{2}I\omega^2 + \frac{1}{2}\frac{(I\omega^2)^2}{kr^2} \qquad [\because I = mr^2]$$

$$= \frac{L^2}{2I} + \frac{1}{2kr^2}(L^2/I)^2 \qquad [\because L = I\omega]$$

$$E = \frac{L^2}{2mr^2} + \frac{L^4}{2kr^2m^2r^4} = \frac{L^2}{2mr^2} + \frac{L^4}{2m^2r^6k}\cdot$$

The quantum restriction that the angular momentum $I\omega$ be quantized according to $\sqrt{J(J+1)}\,\hbar$ will convert this classical result to a quantum-mechanical result.

$$\text{L} = \sqrt{J(J+1)}\,\hbar.$$

The correct allowed energies are,

$$E_J = \frac{\hbar^2}{2mr^2}J(J+1) + \frac{\hbar^4}{2m^2r^6k}J^2(J+1)^2 \qquad ...(4)$$

In Eq. (4), the first term is of major importance. The second is a minor term. It is finally necessary to relate the distorted distance r in the first term to r_0 by means of Eq. (3). In this way, and approximating r by r_0 in the second term, we get

$$E_J = \frac{\hbar^2}{2mr_0^2}J(J+1) - \frac{\hbar^4}{m^2r_0^6k}J^2(J+1)^2 + \frac{\hbar^4}{2m^2r_0^6k}J^2(J+1)^2$$

$$= \frac{\hbar^2}{2mr_0^2}J(J+1) - \frac{\hbar^4}{2m^2r_0^6k}J^2(J+1)^2 \qquad ...(5)$$

We are interested in the wave number

$$\bar{\nu} = \frac{\Delta E}{hc}$$

of the radiation emitted or absorbed as a consequence of changes between energy levels. So we consider energies in these units.

In terms of wave number units,

$$\varepsilon_J = \frac{\hbar}{2\pi Ic}J(J+1) - \frac{\hbar^3m}{4\pi I^3kc}J^2(J+1)^2$$

$$= BJ(J+1) - \frac{4B^3}{\omega^2}J^2(J+1)^2$$

Here, ω is the vibrational energy term $\frac{1}{2\pi c}\sqrt{k/m}.$

Thus, the expression for the rotational energy levels is

$$\varepsilon_J = BJ(J+1) - DJ^2(J+1)^2 \qquad ....(6)$$

Here, D is called the centrifugal distortion constant.

$$D = 4B^3 / \omega^2 \qquad \text{...(7)}$$

D is very small compared with B.

Fig. 23.5 shows the lowering of rotational levels when passing from the rigid to the non-rigid diatomic molecule. The difference between the energy levels increases with increasing value of J.

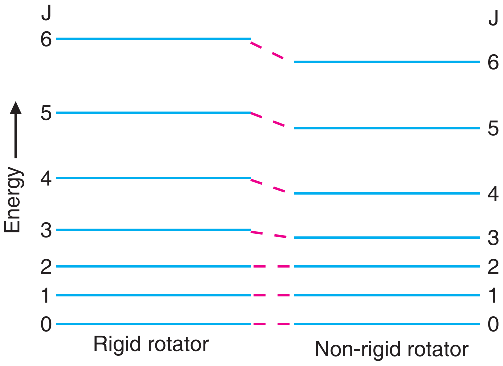


Fig. 23.5

The rotational absorption spectrum is produced due to molecular transition from the state J to the state $(J + 1)$. The selection rule is $\Delta J = +1$.

$$\bar{\nu}_{J+1} - \bar{\nu}_J = \Delta\bar{\nu}$$

$$\Delta\bar{\nu} = B[(J+1)(J+2) - J(J+1)] - D[(J+1)^2(J+2)^2 - J^2(J+1)^2]$$

$$\Delta\bar{\nu} = 2B(J+1) - 4D(J+1)^3\ m^{-1} \qquad \text{...(8)}$$

Thus, we see that the spectrum of a non-rigid diatomic molecule is similar to that of the rigid molecule except that each line is displaced slightly to low frequency (Fig. 23.6).

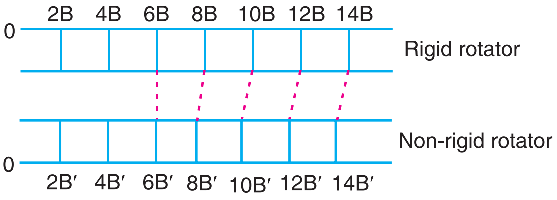


Fig. 23.6

We note from Eq. (8) that the lines are no longer exactly equidistant but their separation decreases slightly with increasing J. The effect, however, is small owing to the smallness of D as compared to B.

Determination of Force Constant from Rotational Spectrum

By making observations on a number of lines in the rotation spectrum of a molecule, and by curve fitting method, the values of B and D are found out.

Once D is known, the J values of lines in the observed spectrum can be determined. Further, knowing B and D, the angular frequency of vibration of the rotating molecule can be calculated using the relation

$$\omega^2 = \frac{4B^3}{D}$$

For hydrogen fluoride, $\omega = 405000\ \text{m}^{-1}$.

The reduced mass of the molecule m is known.

The force constant k is calculated using the equation

$$k = 4\pi^2 c^2 \omega^2 m.$$

For hydrogen fluoride, the force constant is 960 Nm^{-1} which indicates that H-F is a relatively strong bond.

Example 2. *HCl molecule has a rotational constant B value of* 1059.3 m^{-1} *and a centrifugal distortion constant D of* $5.3 \times 10^{-2}\ m^{-1}$. *Estimate the vibrational frequency and force constant of the molecule.*

Sol. $\omega^2 = \dfrac{4B^3}{D}$. Here, ω is the vibrational frequency in wave number unit.

$$\omega = \sqrt{\frac{4B^3}{D}} = \sqrt{\frac{4\times(1059\cdot3)^3}{5\cdot3\times10^{-2}}} = 299516\cdot3\,\text{m}^{-1}.$$

Force constant $k = 4\pi^2 c^2 \omega^2 m$.

$$m = \frac{m_1 m_2}{m_1 + m_2} = \frac{(1\cdot008)(35\cdot46)}{(1\cdot008 + 35\cdot46)}\cdot\frac{1}{6\cdot023\times10^{26}} = 1\cdot627\times10^{-27}\text{ kg}.$$

$$k = 4\pi^2\times(3\times10^8)^2\times(299516.3)^2\times1.627\times10^{-27} = 518.6\text{ Nm}^{-1}.$$

HARMONIC VIBRATIONS OF DIATOMIC MOLECULES

23.10 The Energy of a Diatomic Molecule

Consider a diatomic molecule consisting of two atoms *A* and *B* separated by equilibrium distance r_e. Fig. 23.7 shows how the potential energy of a diatomic molecule varies with the internuclear distance *r*. The minimum of this curve corresponds to the normal configuration of the molecule. The zero of curve is found at $r = r_e$. Any energy in excess of this, for example, ε_1, arises because of extension or compression of the bond. If one atom (*A*) is considered to be stationary on the $r = 0$ axis, the other will oscillate between *B'* and *B"*. If the energy is increased to ε_2, the oscillation will become more vigorous.

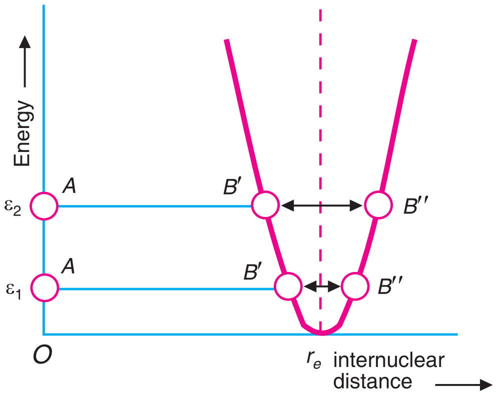


Fig. 23.7

Suppose the bond is distorted from its equilibrium length r_e to new length *r*. The bond, like a spring, obeys Hooke's law. Then,

$$f = -k(r - r_e) \qquad \text{...(1)}$$

Here, *f* is the restoring force, *k* the force constant, and *r* the internuclear distance.

In this case the energy curve is parabolic and has the form

$$E = \frac{1}{2}k(r - r_e)^2 \qquad \text{...(2)}$$

This model of a vibrating diatomic molecule is called "simple harmonic oscillator" model.

23.11 Vibrating Diatomic Molecule as a Harmonic Oscillator

Frequency of Vibrations. In the diatomic molecule we have two bodies of masses m_1 *and* m_2 joined by a spring. Suppose the bond is distorted from its equilibrium length r_e to a new length *r*. Then, restoring forces on each atom of a diatomic molecule are

$$m_1\frac{d^2 r_1}{dt^2} = -k(r - r_e) \qquad \text{...(1)}$$

and

$$m_2\frac{d^2 r_2}{dt^2} = -k(r - r_e) \qquad \text{...(2)}$$

Here, k represents the *force constant.* It is a measure of the stiffness of the bond. r_1 and r_2 are the positions of atoms 1 and 2 relative to the centre of mass of the molecule.

From Art. 19.7, we have,

$$r_1 = \frac{m_2 r}{m_1 + m_2}$$

$$r_2 = \frac{m_1 r}{m_1 + m_2}..$$

Putting the value of r_1 in Eq. (1), we get

$$\frac{m_1 m_2}{m_1 + m_2} \frac{d^2 r}{dt^2} = -k(r - r_e) \qquad \text{...(3)}$$

Since r_e is constant, $\dfrac{d^2 r}{dt^2} = \dfrac{d^2 (r - r_e)}{dt^2}$

$$\therefore \qquad \frac{m_1 m_2}{m_1 + m_2} \frac{d^2 (r - r_e)}{dt^2} = -k(r - r_e) \qquad \text{...(4)}$$

Put $r - r_e = x$, and $\dfrac{m_1 m_2}{m_1 + m_2} = \mu.$

Here, x represents the displacement of the bond length from its equilibrium position.

Eq. (4) becomes

$$\mu \frac{d^2 x}{dt^2} = -kx$$

or
$$\frac{d^2 x}{dt^2} + \frac{k}{\mu} x = 0$$

or
$$\frac{d^2 x}{dt^2} + \omega^2 x = 0$$

where $\omega^2 = k/\mu$.

This is the equation of simple harmonic motion with frequency of vibration,

$$\nu = \frac{1}{2\pi} \sqrt{\left(\frac{k}{\mu}\right)}.$$

In m^{-1} unit (wave number),

$$\bar{\nu} = \frac{1}{2\pi c} \sqrt{\left(\frac{k}{\mu}\right)}.$$

Energy levels. Vibrational energies, like other molecular energies are quantised. The permitted vibrational energies for any particular system can be calculated from Schrodinger equation. The eigen values for the energy (in joules) of a linear harmonic oscillator are given by

$$E_v = \left(v + \frac{1}{2}\right) h\nu.$$

Here v is the *vibrational quantum number*, equal to zero or an integer.

Converting to the spectroscopic units, m^{-1}, we have,

$$\varepsilon_v = \frac{E_v}{hc} = \left(v + \frac{1}{2}\right) \bar{\nu}.$$

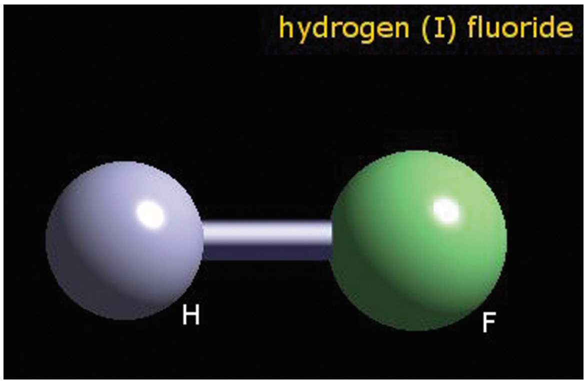
Here, ε_v is called *term value*. Its units is m^{-1}.

$\bar{\nu}$ is the vibrational frequency of the oscillator in wave numbers.

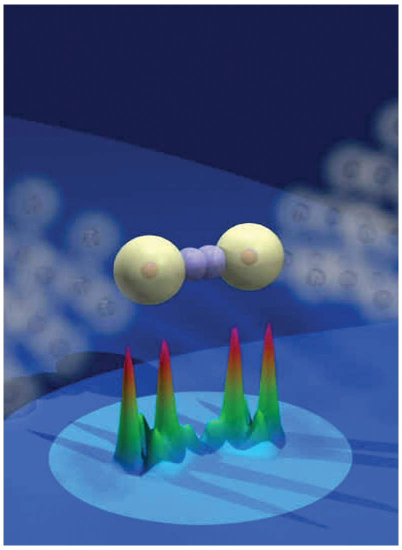
Zero-point energy in $m^{-1} = \frac{1}{2}\bar{\nu}$.

The *selection rule* for the harmonic oscillator undergoing vibrational changes is

$$\Delta v = \pm 1.$$



HF (Hetronuclear Diatomic Molecules)



H_2(Homonuclear Diatomic Molecules)

Vibrational energy changes will only give rise to an observable spectrum if the vibration can interact with radiation, *i.e.*, if the vibration involves a change in the dipole moment of the molecule. Thus, vibrational spectra will be observable only in heteronuclear diatomic molecules (like HF, HCl, HBr) since homonuclear molecules (like H_2, N_2 and O_2) have no dipole moment. Applying the selection rule, we have

$$\varepsilon_{v+1 \to v} = (v+1+1/2)\bar{\nu} - (v+1/2)\bar{\nu} = \bar{\nu}$$

for emission and

$$\varepsilon_{v \to v+1} = \bar{\nu}$$

for absorption, whatever the initial value of v.

It means that all the vibrational lines obtained from harmonic oscillator are of the same frequency. The allowed vibrational energy levels and transitions between them for a diatomic molecule undergoing simple harmonic motion are shown in Fig. 23.8.

Pure vibrational spectra are observed only in liquids. This is because interactions between neighbouring molecules prevent their rotational motion.

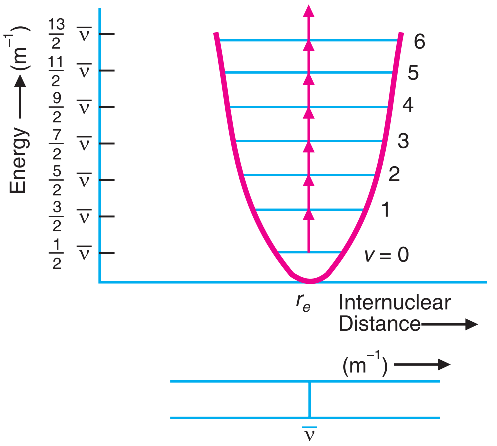


Fig. 23.8

EXAMPLE 3. *In the near infra-red spectrum of HCl molecule there is single intense band at 2885.9 cm^{-1}. Assuming that it is due to the transition between vibrational levels, show that the force constant k is 480 Nm^{-1}. (Given : M_H = 1.68 × 10^{-27} kg).* **(Meerut 77, 86)**

SOL. $\bar{\nu} = 2885.9 \text{ cm}^{-1} = 288590 \text{ m}^{-1}$.

Now, $\bar{\nu} = 1/\lambda = \nu/c$.

Frequency of vibration is

$$\nu = \bar{\nu} c = 288590 \times (3 \times 10^{8}) = 8 \cdot 6577 \times 10^{13} \text{ Hz.}$$

Now, $\nu = \frac{1}{2\pi}\sqrt{\left(\frac{k}{\mu}\right)}$, where k is the force constant and μ is the reduced mass of the molecule.

Therefore,

$$k = 4\pi^{2} \mu \nu^{2}.$$

The reduced mass of $H^{1}\, Cl^{35}$ is given by

$$\mu = \frac{M_H M_{Cl}}{M_H + M_{Cl}} = \frac{1 \times 35}{1+35} M_H = \frac{35}{36} \times (1 \cdot 68 \times 10^{-27} \text{ kg}) = 1 \cdot 63 \times 10^{-27} \text{ kg}$$

$$\therefore \quad k = 4 \times (3.14)^{2} \times (1.63 \times 10^{-27})\, (8.6577 \times 10^{13})^{2} = 482.3 \text{ Nm}^{-1}.$$

Example 4. *The force constant of the bond in CO molecule is 187 Nm*$^{-1}$*. Calculate the frequency of vibration of the molecule and the spacing between its vibrational energy levels in eV. Given that reduced mass of CO* = 1.14 × 10^{-26} *kg ,* h = 6.6 × 10^{-34} *Js and* 1*eV* = 1.60 × 10^{-19} *J.*

Sol. The frequency of vibration of the molecule is

$$\nu_{osc} = \frac{1}{2\pi}\sqrt{\left(\frac{k}{\mu}\right)} = \frac{1}{2 \times 3 \cdot 14}\sqrt{\left(\frac{187 \text{ Nm}^{-1}}{1 \cdot 14 \times 10^{-26} \text{ kg}}\right)} = 2 \cdot 04 \times 10^{13} \text{ Hz}$$

The vibrational energy of a diatomic molecule is

$$E_v = h\nu\, (v + 1/2),\ v = 0, 1, 2 \ldots$$

The separation between two successive vibrational energy levels is

$$\Delta E = E_{v+1} - E_v = h\nu\left(v + \frac{3}{2}\right) - h\nu\left(v + \frac{1}{2}\right) = h\nu$$

$$= (6.63 \times 10^{-34}) \times (2.04 \times 10^{13}) \text{ J} = 8.44 \times 10^{-2} \text{ eV.}$$

Example 5. *Given that the spacing between the vibrational levels of CO molecule is* 8.45 × 10^{-2} *eV of energy. Find the force constant of the molecule.*

Sol.

$$\Delta E = 8.45 \times 10^{-2} \text{ eV} = 8.45 \times 10^{-2} \times 1.602 \times 10^{-19} \text{ J}$$

$$\therefore \quad \nu = \frac{\Delta E}{h} = \frac{8 \cdot 45 \times 10^{-2} \times 1 \cdot 602 \times 10^{-19}}{6 \cdot 63 \times 10^{-34}} = 2 \cdot 04 \times 10^{13} \text{ hertz.}$$

The reduced mass of CO molecule = μ = 1.14 × 10^{-26} kg.

$$k = 4\pi^{2} \nu^{2} \mu = 4 \times (3.14)^{2} \times (2.04 \times 10^{13})^{2} \times 1.14 \times 10^{-26}$$

$$= 187 \text{ Nm}^{-1}.$$

Example 6. *Calculate the vibrational energy levels of an HCl molecule, assuming the force constant to be 516 Nm*$^{-1}$.

Sol.

$$\mu = (35 / 36) \times 1.0078 \times 1.66 \times 10^{-27} \text{ kg.}$$

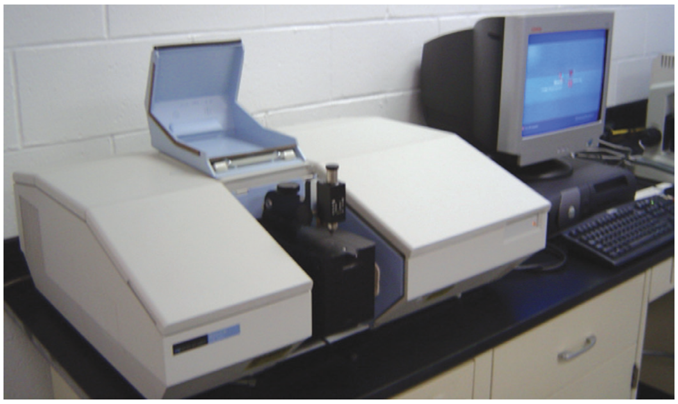
$$\nu_0 = \frac{1}{2\pi}\left(\frac{516}{\mu}\right)^{1/2}$$

The energy levels are then given by

$$h\nu_0\, (v + 1/2) = 6 \cdot 6 \times 10^{-34} \times \frac{1}{2\pi}\left(\frac{516 \times 36}{35 \times 1 \cdot 0078 \times 1 \cdot 66 \times 10^{-27}}\right)^{\frac{1}{2}} (v + 1/2) \text{ J}$$

$$\sim\ 5.9 \times 10^{-20}(v + 1/2)\text{J}$$
$$\sim\ 0.37\ (v + 1/2)\ \text{eV}\ (v = 0, 1, 2, ...)$$

23.12 Infra-red Spectra



Infra-red Spectrum Equipment.

Suppose a vibrational transition takes place within a given electronic level. Then the wavelength of the emitted radiation would be of the order of 12,000 nm which would be in the near infra-red region of the spectrum. The infra-red (I. R.) radiation absorbed by a pure liquid or by a material in solution can be studied and *well-defined* absorption lines can be recorded. This is because, in the liquid phase, no well-defined rotational energy levels exist (interaction between adjacent molecules inhibit rotation). So rotational structure is not found on a vibrational absorption band spectrum.

However, the I.R. spectrum of a simple (diatomic) molecule in the gas phase, when examined with high resolution instrument, shows rotational component lines in addition to vibrational lines. When vibrational energy changes, it is accompanied by rotational energy change. When this happens, each absorption band (which is due to a particular vibrational energy change) is found to consist of a number of relatively closely spaced absorption lines. They appear as a broad streak, called the vibration-rotation band. The rotational components of the band can be related to the rotational energy changes that occur simultaneously with the vibrational energy change. A study of such vibration-rotation bands (I.R. spectrum) enables us to find force constant, moment of inertia and equilibrium bond length data of the molecule concerned.

A pure rotational transition without any change in either the vibrational or the electronic energy is possible. The resulting spectral lines will be in the far infra-red. These lines constitute *rotational* spectra.

23.13 Theory of the Origin of the Vibration-Rotation Spectrum of a Molecule

(See Section 19.8)

23.14 The Anharmonic Oscillator

Real molecules do not obey exactly the laws of simple harmonic motion. Real molecules dissociate at large values of r and acquire very large amounts of energy as r decreases. Fig. 23.9 shows the energy curve for a typical diatomic molecule undergoing anharmonic extensions and compressions together with (dotted) the ideal simple harmonic parabola.

A purely empirical expression which fits this curve to a good approximation was derived by P. M. Morse, and is called the Morse function :

$$E = D_e\,[1 - \exp.\ \{a\,(r_e - r)\}]^2 \qquad ...(1)$$

where D_e is the dissociation energy, a is constant for a particular molecule, r_e is the value of internuclear distance r that corresponds to the minimum of Morse curve.

Energy Levels. When Eq. (1) is used in the Schrodinger equation, the allowed vibrational energy levels are found to be

$$\varepsilon_v = (v+1/2)\bar{\nu}_e - (v+1/2)^2\, \bar{\nu}_e\, x_e\, m^{-1}\ (v = 0,1,2,...) \qquad ...(2)$$

Here, $\bar{\nu}_e$ is called the *equilibrium oscillation frequency* of the molecule expressed in wave numbers and is the value for small displacements. x_e is the corresponding anharmonicity constant. $x_e \approx 0.01$.

The selection rules for the anharmonic oscillator are found to be $\Delta v = \pm 1, \pm 2, \pm 3, ...$

Normally only the lines of $\Delta\ v = \pm 1, \pm 2$ and ± 3, have observable intensity.

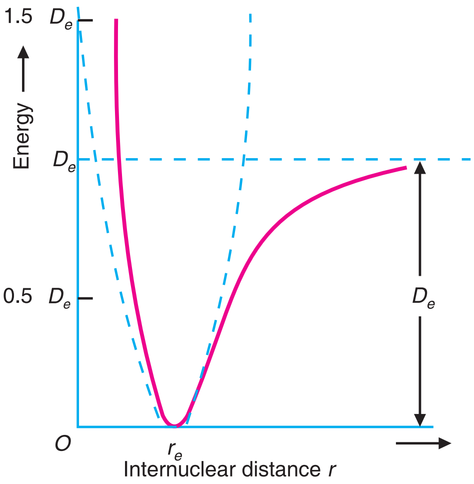


Fig. 23.9

Vibration bands are almost invariably observed in absorption. At room temperature nearly all the molecules in a particular sample have only the zero point vibrational energy and exist in $v = 0$ state. The absorption of radiation must, therefore, result in transitions starting from $v = 0$.

1. The *fundamental* absorption band corresponds to

$$v = 0 \rightarrow v = 1,\ \Delta\ v = +\ 1, \text{ with considerable intensity.}$$

$$\begin{aligned}\Delta\ \varepsilon &= \varepsilon_{v=1} - \varepsilon_{v=0}\\ &= (1+1/2)\bar{\nu}_e - x_e\,(1+1/2)^2\,\bar{\nu}_e - \{1/2\,\bar{\nu}_e - (1/2)^2\, x_e\,\bar{\nu}_e\}\\ &= \bar{\nu}_e\,(1-2x_e)\, m^{-1} \qquad ...(4)\end{aligned}$$

2. Band corresponding to $v = 0 \rightarrow v = 2$, $\Delta\ v = +\ 2$, with small intensity, is called *first overtone.*

$$\begin{aligned}\Delta\ \varepsilon &= (2+1/2)\bar{\nu}_e - x_e\,(2+1/2)^2\,\bar{\nu}_e - \{1/2\,\bar{\nu}_e - (1/2)^2\, x_e\,\bar{\nu}_e\}\\ &= 2\bar{\nu}_e\,(1-3x_e)\,\mathrm{m}^{-1} \qquad ...(5)\end{aligned}$$

3. Band corresponding to $v = 0 \rightarrow v = 3$, $\Delta\ v = +\ 3$, with normally negligible intensity is called *second overtone.*

$$\begin{aligned}\Delta\ \varepsilon &= (3+1/2)\bar{\nu}_e - (3+1/2)^2\,\bar{\nu}_e x_e - \{1/2\,\bar{\nu}_e - (1/2)^2\, x_e\,\bar{\nu}_e\}\\ &= 3\bar{\nu}_e\,(1-4x_e)\,\mathrm{m}^{-1} \qquad ...(6)\end{aligned}$$

The allowed vibrational energy levels and these three absorption transitions are shown in Fig. 23.10.

4. If the temperature is raised, the population of the $v = 1$ state may become appreciable. The transitions from energy levels with the selection rule $\Delta\ v = 1$ are called *hot bands*. The first hot band corresponds to

$$v = 1 \rightarrow v = 2,\ \Delta\ v = +1, \text{ normally very weak.}$$

$$\begin{aligned}\Delta\varepsilon &= 2\frac{1}{2}\bar{\nu}_e - 6\frac{1}{4}x_e\,\bar{\nu}_e - \left\{1\frac{1}{2}\bar{\nu}_e - 2\frac{1}{4}x_e\,\bar{\nu}_e\right\}\\ &= \bar{\nu}_e\,(1-4x_e)\,\mathrm{m}^{-1}\end{aligned}$$

Their nature may be confirmed by raising the temperature of the sample when a true hot band will increase in intensity.

Example 7. *The fundamental band for CO is centred at 2143.3 cm^{-1} and first overtone at 4259.7 cm^{-1}. Calculate* $\bar{\nu}_e$ *and* $\bar{\nu}_e x_e$.

Sol. We know that frequencies of fundamental and first overtone bands are given by

$$\bar{\nu}_1 = (1-2x_e)\bar{\nu}_e$$

$$\bar{\nu}_2 = (1-3x_e)2\bar{\nu}_e$$

and

$$(1-2x_e)\bar{\nu}_e = 214330\,\text{m}^{-1} \quad \ldots(1)$$

$$(1-3x_e)2\bar{\nu}_e = 425970\,\text{m}^{-1} \quad \ldots(2)$$

Solving (1) and (2),

$$x_e = 0.006,\ \bar{\nu}_e = 224992\ \text{m}^{-1}.$$

$$\bar{\nu}_e x_e = 1350\ \text{m}^{-1}.$$

Example 8. *The spectrum of HCl shows a very intense absorption at 2886 cm^{-1}, a weaker one at 5668 cm^{-1}, and very weak one at 8347 cm^{-1}. Find the equilibrium frequency of the molecule and the anharmonicity constant.*

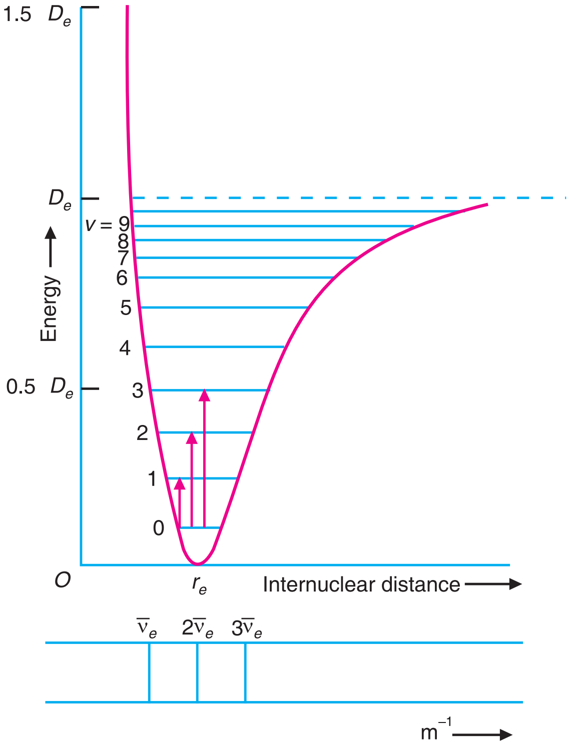


Fig. 23.10

Sol. We solve any two of the following three equations :

$$\bar{\nu}_e = (1-2x_e) = 2886$$

$$2\bar{\nu}_e\ (1-3x_e) = 5668$$

$$3\bar{\nu}_e\ (1-4x_e) = 8347\ \text{cm}^{-1}$$

$$\bar{\nu}_e = 2990\ \text{cm}^{-1}, \qquad x_e = 0.0174.$$

The force constant of the bond in HCl may be calculated by using the formula,

$$k = 4\pi^2\bar{\nu}_e^2 c^2 \mu\,\text{Nm}^{-1}.$$

23.15 Fundamental and Overtone Frequencies

The frequency corresponding to $v = 0$ to $v = 1$ transition is called the *fundamental frequency*. We sometimes observe frequencies due to $v = 0$ to $v = 2$ or even $v = 0$ to $v = 3$ transitions. Such frequencies are called *overtone frequencies*. Since the frequencies of first and second overtone bands are approximately 2 to 3 times the frequency of the fundamental band, they appear in regions of the shorter wavelengths than do the fundamental.

23.16 Calculation of Force Constant

The vibrational frequency of fundamental = $\bar{\nu}_e(1-2x_e)\,\text{m}^{-1}$

The vibrational frequency of the first overtone = $2\bar{\nu}_e\,(1-3x_e)\,\mathrm{m}^{-1}$

The vibrational frequency of the second overtone = $3\bar{\nu}_e\,(1-4x_e)\,\mathrm{m}^{-1}$.

By studying the vibration spectra of a diatomic molecule, the above frequencies could be identified. Knowing these frequencies, we can slove any two of the above three equations and calculate the equilibrium vibrational frequency $\bar{\nu}_e$. The reduced mass of the molecule (μ) is known.

The force constant (k) of the bond in the diatomic molecule may be calculated from the equation,

$$k = 4\pi^2 c^2 \mu \bar{\nu}_e^2 .$$

EXAMPLE 9. *The fundamental and the first overtone frequencies of the diatomic molecule HCl behaving as anharmonic oscillator are found to be* 2886 *cm*$^{-1}$ *and* 5668 *cm*$^{-1}$ *respectively. If the reduced mass of HCl be* 1.63×10^{-27} *kg, calculate the force constant.*

SOL.

$$\bar{\nu}_e\,(1-2x_e) = 2886 \qquad \text{...(1)}$$

$$2\bar{\nu}_e\,(1-3x_e) = 5668 \qquad \text{...(2)}$$

Solving Eqs. (1) and (2),

$$\bar{\nu}_e = 2990\ \mathrm{cm}^{-1} = 2990 \times 10^2\ \mathrm{m}^{-1}$$

$$\begin{aligned} k &= 4\pi^2 c^2 \mu \bar{\nu}_e^2 \\ &= 4 \times (3.14)^2 \times (3 \times 10^8)^2 \times 1.63 \times 10^{-27} \times (2990 \times 10^2)^2 \\ &= 516\ \mathrm{Nm}^{-1}. \end{aligned}$$

EXERCISE

1. Describe the main features of molecular spectra.
2. Derive an expression for energy of rotational states of a non-rigid rotator and show how force constant of the molecule can be determined.
3. Give the elementary theory of the origin of the vibrational spectrum of a molecule.
4. Discuss how the study of vibrational spectrum of a diatomic molecule enables us to determine the anharmonicity constant and equilibrium frequency of vibration.
5. Explain, giving theory, the method of finding the force constant of a diatomic molecule from its anharmonic vibration spectrum.

CHAPTER 24

SPECTROSCOPIC TECHNIQUES

AT A GLANCE

24.1 Introduction
24.2 Constant Deviation Spectrograph
24.3 Recording the Spectrum
24.4 Enhancement of Spectra: Computer Averaging
24.5 Ultraviolet Spectrascopy
24.6 Quartz Spectrograph for Near U.V. Region
24.7 Littrow Spectrograph
24.8 Concave Grating Vacuum Spectrograph
24.9 Infrared Spectroscopy
24.10 Wadsworth Prism-Mirror Spectrograph
24.11 Grating Spectrograph for Infra-red
24.12 Absorption Spectroscopy
24.13 Block Diagram of an Absorption Spectrometer
24.14 Principle of a Double-Beam Spectrophotometer
24.15 Double-beam Infrared Spectrometer
24.16 Complementary Character of Raman and Infra-red Spectra

24.1 Introduction

The ranges of electromagnetic radiation, expressed in the units usually used for Infrared, Visible and Ultraviolet regions, are shown in Table 1.

TABLE 1

	Infrared	*Visible*	*Ultraviolet*
$\lambda(m)$	0.00003–0.0000025	7×10^{-7}–3×10^{-7}	3×10^{-7}–1×10^{-7}
$\lambda(mn)$		700–300	300–100
$\overline{v}$	30000–400000 m^{-1}		

SPECTROSCOPY OF THE VISIBLE REGION

24.2 Constant Deviation Spectrograph

In a spectrograph, the spectra are photographed. In constant deviation spectrograph, the telescope (*T*) and the collimator (*C*) are fixed at right angles to one another [Fig. 24.1]. A constant deviation prism (*P*) is mounted on the prism table. The prism table can be rotated by means of a micrometer screw carrying a drum (*D*) at one end. The drum is calibrated in terms of wavelengths.

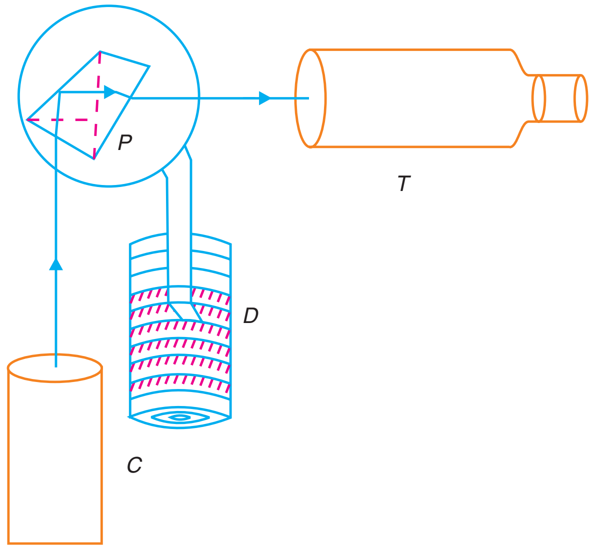


Fig. 24.1

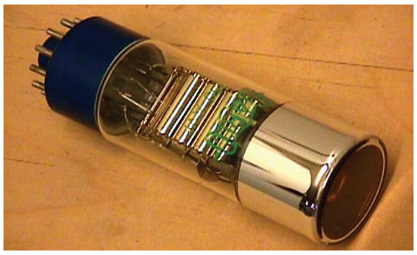
Overall Size of Spectrograph Compared to the Telescope.

When the table is rotated, the lines in the spectrum of source of light appear one by one on the intersection of the cross wires of the telescope. The wavelength of any one line can then be read off from the drum. This instrument is called the Hilger wavelength prism spectrometer. For taking photograph of the lines, the eyepiece in the telescope is replaced by a camera.

In order to set the instrument for work, the slit is illuminated with light of known wavelength. The drum is set so that the correct wavelength is indicated on the drum. The prism is placed on the prism table and its position is adjusted till the line is on the intersection of the crosswires. The prism is clamped tight to the table in this position. Any other wavelength can be determined by rotating the drum till the spectral line is on the crosswire and taking the reading of the drum.

24.3 Recording the Spectrum

In modern spectrometers, the detector is rarely the simple photographic plate. One of the most sensitive and useful devices in the visible and ultra-violet region is the photomultiplier tube.



A photomultiplier Tube.

The photomultiplier tube consists of a light-sensitive surface which emits electrons when light falls upon it. The tiny electron current is amplified and applied to a pen-recorder. Fig. 24.2 shows a spectrometer employing a photomultiplier as detector and recording the spectrum graphically on chart paper. The sensitive element of the photomultiplier is situated at the point *A*. Rays have been drawn to show the points *A* and *B* at which two frequencies, v_1 and v_2, are focused. The physical width of the beam falling on the detector can be limited by the provision of an 'exit slit' just in front of the detector entrance.

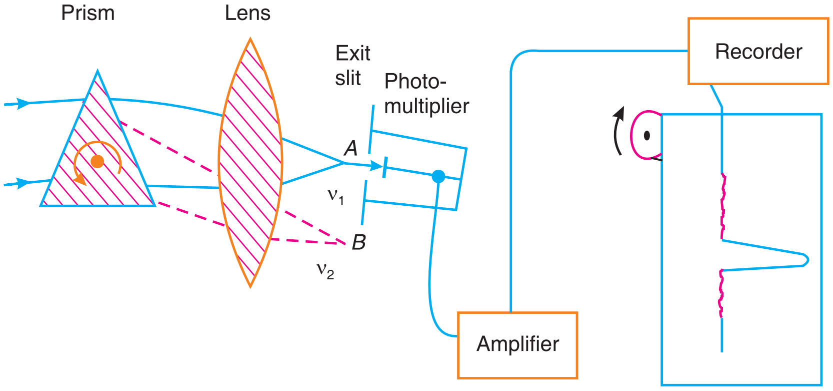


Fig. 24.2

The frequency of the light falling on the photomultiplier is altered by steady rotation of the prism. Thus the spectrum is *scanned* between the frequencies represented by *A* and *B*. Such a picture is called a spectrum in the *frequency domain*, to indicate that it records the detector output against frequency.

24.4 Enhancement of Spectra: Computer Averaging

All modern spectrometers use some form of electronic amplification to magnify the signal produced by the detector. So every recorded spectrum has a background of random fluctuations caused by spurious electronic signals produced by the detector, or generated in the amplifying equipment. These fluctuations are usually referred to as 'noise'. A signal-to-noise (S/N) ratio of 3 or 4 is necessary for unambiguous recognition of a signal.

Signal-to-noise ratio can be improved by using a computer-averaging technique. It is called '*computer averaging of transients*' or the CAT technique. This technique involves recording the spectrum stepwise into a computer. We sample a frequency domain spectrum at, say, 2000 closely spaced points. We store the intensity at each point in 2000 separate computer memory locations. We repeat this process as many times as we wish, but each time *adding* the new data into that already existing. In any one scan a weak signal may not be visible above the noise level. After n summed scans, the signal will be n times larger in the store. But the noise is random. It will sometimes contribute to the store in a positive sense and sometimes negatively. So the noise will accumulate less rapidly. It may be shown that n scans increase the noise level in the store by $n^{1/2}$. So the net gain in S/N is $n/n^{1/2} = n^{1/2}$.

ULTRAVIOLET SPECTROSCOPY

24.5 Introduction

The ultraviolet region of the spectrum extends from 300 nm to about 10 nm.

Sources. An electric arc of carbon, iron or of other materials, mercury vapour lamps and discharge of electricity through hydrogen contained in quartz tubes are used in the laboratory as sources of U.V. light. Mercury arc enclosed in quartz tube is generally used for medical purposes.

For extremely short waves, vacuum spark discharge should be employed.

Production of spectrum. The ultra-violet spectrum is produced either by a *prism* or by a *grating*.

In the **prism method,** the material of the prism should be such that it does not absorb the radiations to be studied. The following materials are suitable for ranges shown against each.

TABLE 2. Materials and ranges of UV-radiation

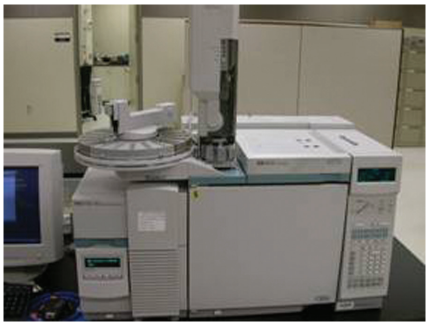
Materials		*Range nm*
Glass	up to	330
UV-glass	...	200
Quartz	...	85
Fluorite	...	120
SiF crystal (artificial)	...	110

The associated lenses should also be of the same material.

In the **grating method**, the range is extended without limit by the use of reflection gratings, concave or plane. Grating is the *only method* for the production of UV-spectrum below 110 nm.

Recording of spectrum. In order to record the UV-spectra, the photographic method, using special plates such as Ilford Q-plates, is generally employed.

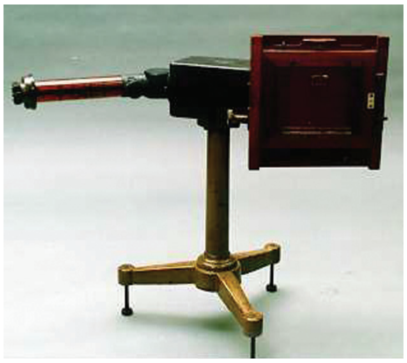
All wavelengths are determined by photographing the standard wavelengths touching along one edge of the spectrum of unknown wavelengths, and by interpolation the unknown wavelength is computed using a *comparator*.



Ultraviolet Sepctroscopy.

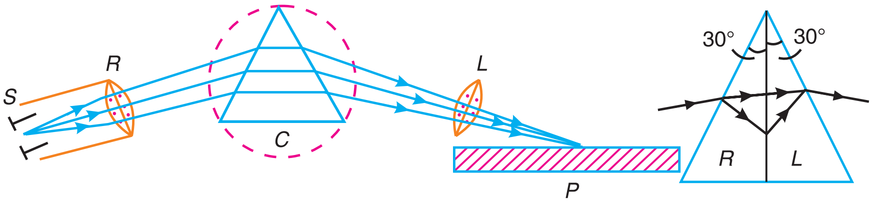
24.6 Quartz Spectrograph for Near U.V. Region

Glass absorbs U.V., radiations and cannot be used for work in the ultraviolet region. Hence quartz lenses and prism are used. Fig. 24.3 (a) shows the optical components of a quartz spectrograph. It has a collimator with an adjustable slit and a right-handed quartz collimating lens R. The camera consists of a left-handed quartz objective L and photographic plate P. In order to get the whole spectrum in focus on the photographic plate, the plate is inclined at 20°



Quartz Spectrograph.

to the axis of the telescope. A Cornu prism (*C*) is mounted in the minimum deviation position on the prism table between the collimator and camera tubes. The Cornu prism [24.3 (*b*)] consists of two 30° right angled prisms cemented together, one being made of right handed quartz and the other of left-handed quartz. When used at minimum deviation position, the prism gives a single emergent beam because the phase difference introduced in one half is removed in the other.



(*a*) (*b*)

Fig. 24.3

Light from the slit (*S*) is rendered parallel by the lens *R*. After refraction through the Cornu prism, light emerges out in a direction symmetrical with the incident beam and enters the camera. The camera lens *L* brings the spectrum to a perfect focus on the photographic plate *P*. The spectrum is focussed using a fluorescent screen. The screen is then removed and the photographic plate put in position. The U.V., spectrum is thus photographed. Wavelengths may be computed with a comparator.

24.7 Littrow Spectrograph

Fig. 24.4 shows a *Littrow-spectrograph*. *M* is a front-silvered plane mirror. *L* is a lens of fused quartz (silica). *P* is a 30°–fused quartz prism backed by a mirror *H*. *S* is the slit. *PP* is the photographic plate. These are all encased in a vacuum chamber.

The Littrow arrangement is compact.

Littrow Spectrograph.

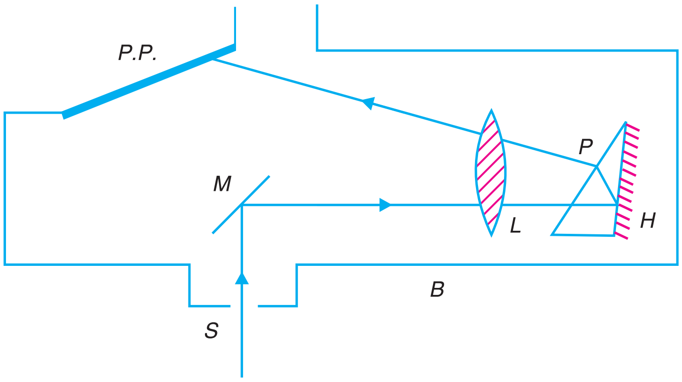


Fig. 24.4

24.8 Concave Grating Vacuum Spectrograph

The region below 170 nm is called the vacuum ultraviolet region. Fig. 24.5 shows a Concave grating vacuum spectrograph with Eagle's mounting. It is used up to 30 nm. *V* is the vacuum chamber, S_2 the source, S_1 the slit, *PP* the photographic plate and *G* the concave grating. *PP* and *G* are on the Rowland circle. The light from the slit falls on the grating *G* and the diffraction spectrum is received on the photographic plate.

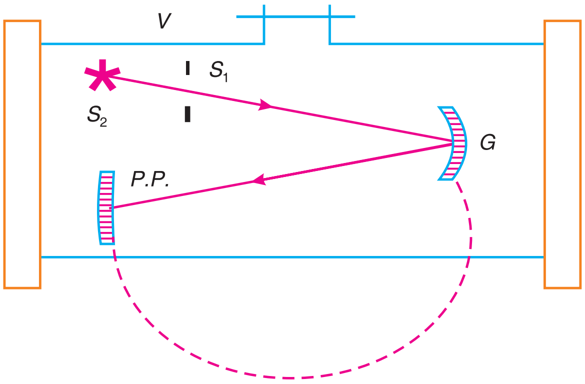


Fig. 24.5

INFRARED SPECTROSCOPY

24.9 Introduction

The infra-red region of the spectrum extends beyond the red end of the visible spectrum (720 nm) up to 400 microns (1 micron = 1μ = 10^{-4} cm).

Sources : *Nernst glower* and *Globar* are the most common sources of infrared radiations. The Nernst glower consists of a filament prepared from a mixture of rare earths-particularly the oxides of thorium, zirconium and cesium. The glower is maintained at incandescence electrically. The globar consists of a rod of silicon carbide. It is also electrically heated.

Dispersing systems. The spectrographs used for studying infrared spectra employ mirrors instead of lenses both before and after dispersion. The prism materials depend upon the range of wavelengths to be studied. No single material is transparent to the entire region. The following table shows the wavelength range over which the prism materials are reasonably transparent.

Infrared Spectroscopy Equipment.

Material	*Wavelength Range*
Glass	Visible to 2.5μ
Quartz	Visible to 3.5μ
Fluorite	3μ to 9μ
Rock salt (Sodium chloride)	8μ to 16μ
Sylvine (K Cl)	15μ to 20μ
Potassium Bromide (K Br)	18μ to 28μ
Thallium Bromide	25μ to 40μ

24.10 Wadsworth Prism-Mirror Spectrograph

S is the source and the radiations from it are focussed on the slit S_1 by the concave mirror M_1 (Fig. 24.6).

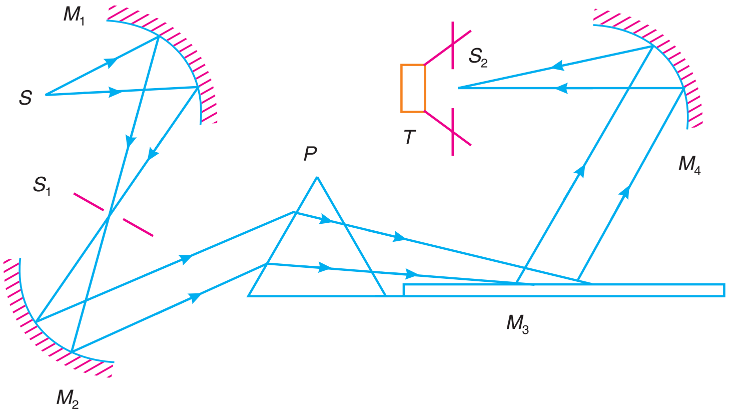


Fig. 24.6

The light from S_1 is rendered parallel by the concave mirror M_2. The parallel beam passes through the prism-mirror combination and after emergence, falls on the concave mirror M_4 and brought to a focus on the exit slit S_2. A linear thermopile *T* is kept at S_2. The prism is rotated so that spectral lines of different wavelengths pass through *T* in succession. The mirrors and prism are mounted inside a box to avoid stray light. The spectrograph is set focussed for visible monochromatic light and holds good for all other wavelengths.

24.11 Grating Spectrograph for Infra-red

Fig. 24.7 shows a grating spectrograph for studying IR-radiation.

The infra-red radiations are focussed on a small hole at the centre of a plane metallic mirror M_1. The hole is at the focus of the concave mirror C_1M_1. The subsequent paths are shown in Fig. *G* is the grating. CM_2 and CM_3 are two concave mirrors between which is placed another plane mirror. M_2 is the exit slit. The beam is finally focussed on the thermopile *T*.

The grating is turned by a graduated drum to *scan* the spectrum, *i.e.*, to bring different wavelength regions at the exit slit. From this rotation, the diffraction angles corresponding to the different orders can be obtained and the wavelength λ can be evaluated.

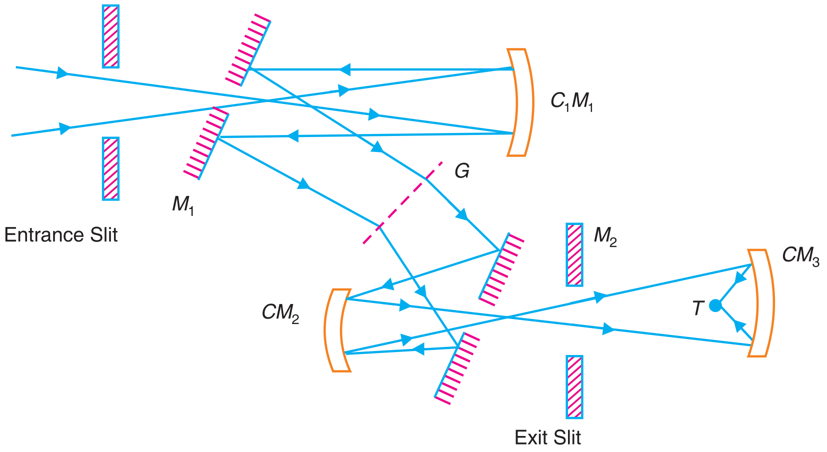


Fig. 24.7

24.12 Absorption Spectroscopy

Light from a source, which emits radiations of all wavelengths is passed through the sample under study. The transmitted light is spectrally studied to find the absorption at different wavelengths.

Fig. 24.8 shows the setting for recording the absorption spectrum in the visible region.

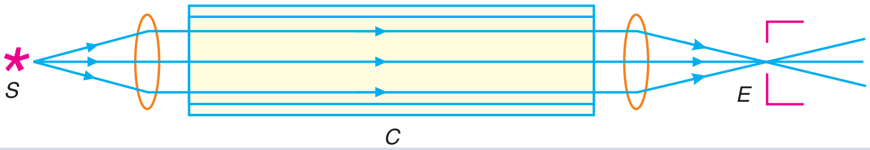
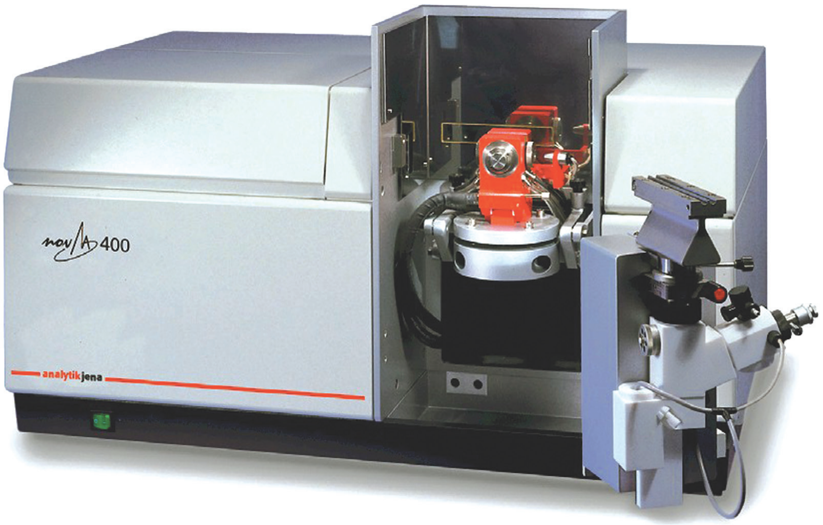


Fig. 24.8

S is a filament source giving a continuous spectrum.

C is a cell containing the sample under test.

E is the entrance slit of a spectrograph.


novAA 400
analytikjena

Absorption Spectroscopy.

24.13 Block Diagram of an Absorption Spectrometer

Fig. 24.9 shows the block diagram of an absorption spectrometer for use in the infra-red, visible, and ultra-violet regions.

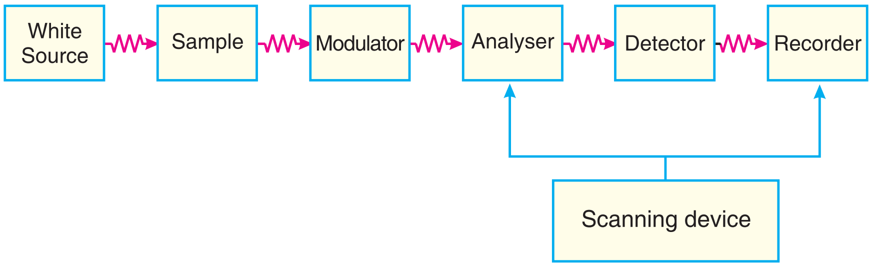

White Source
Sample
Modulator
Analyser
Detector
Recorder
Scanning device

Fig. 24.9

The radiation from a *white* source is directed by some guiding device (lens) on to the sample. From the sample, the radiation passes through an analyser (prism). The analyser selects the frequency reaching the detector at any given time. The signal from the detector passes to a recorder which is synchronized with the analyser so as to produce a trace of the absorbance as the frequency varies.

Caution!
High magnetic field
Attention!
Champ magnétique intense
Achtung!
Starkes Magnetfeld
MAGNET
POWER
LOCK
PERKIN-ELMER Zeeman 5100
PERKIN ELMER AS-60/70
PERKIN ELMER

Absorption Spectrometer.

24.14 Principle of a Double-Beam Spectrophotometer

In most absorption studies, the *double-beam principle* is used.

Fig. 24.10 shows the schematic diagram of a double-beam spectrophotometer.

The incident beam is split at the plate P_1 into two beams A and B. The two beams are energetically and optically identical. The beam A passes through the absorption sample. The beam B passes through a variable intensity control R. The beams A' and B' are then to be balanced at each wavelength λ.

The percentage cut-off of intensity needed at R to secure balance gives the percentage absorption by the sample in path A.

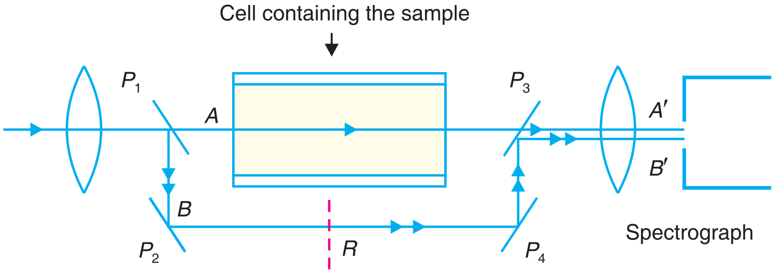

Cell containing the sample
P1
A
B
P2
R
P3
P4
A′
B′
Spectrograph

Fig. 24.10

Advantage. The double beam system makes the working independent of time variations in source intensity.

24.15 Double-beam Infrared Spectrometer

Fig. 24.11 shows the block diagram of a double-beam infrared spectrometer.

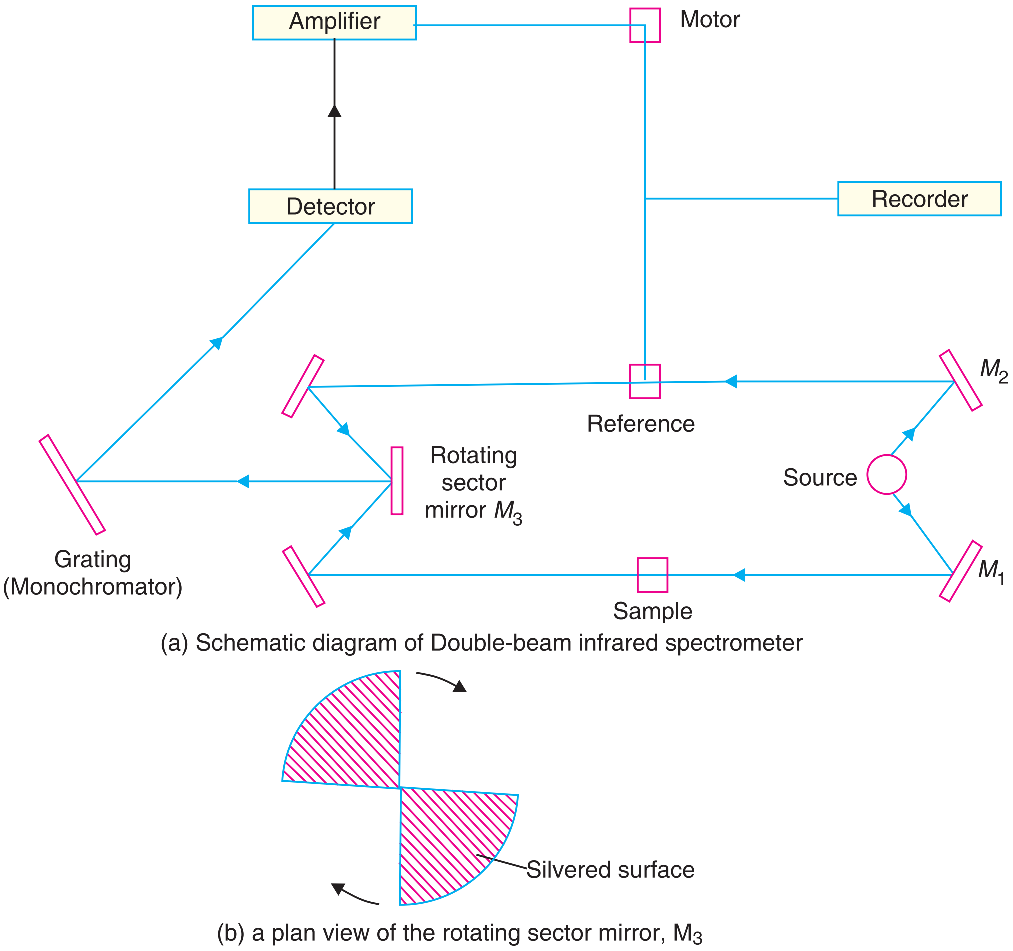

Amplifier
Motor
Detector
Recorder
M_2
Reference
Rotating sector mirror M_3
Source
Grating (Monochromator)
M_1
Sample
(a) Schematic diagram of Double-beam infrared spectrometer
Silvered surface
(b) a plan view of the rotating sector mirror, M3

Fig. 24.11

Double-beam Infrared Spectrometer.

The essential components of the spectrometer are:

(*i*) source of radiation (*ii*) fore optics (*iii*) monochromator
(*iv*) detector with an amplifier and (v) recorder.

Principle. The source radiation is divided into two by means of the mirrors M_1 and M_2. One beam is brought to a focus at the sample space, while the other follows an exactly equivalent path and is referred to as the reference beam. The two beams meet at the sector mirror M_3. As this mirror rotates it alternately reflects the reference beam, or allows the sample beam through the spaces, into the monochromator. Thus the detector 'sees' the sample beam and reference beam alternately.

If a sample, capable of absorbing energy from the beam at the particular frequency passed by the monochromator, is now placed in the sample beam, the detector will receive a signal alternating in intensity, since the sample beam carries less energy than the reference beam. This alternating signal is amplified electronically. A calibrated attenuator is driven into the reference beam until the signal is reduced to zero, i.e., until sample and reference beams are again balanced. The distance moved by the attenuator is a direct measure of amount of the energy absorbed by the sample.

Working (*i*) **Source of radiation.** The source is always some form of filament which is maintained at red-or white-heat by an electric current. Two common sources are the Nernst filament and the 'globar' filament. The Nernst glower consists of a spindle of rare earth oxides about 2.5 cm long and 0.25cm in diameter. It has very good emissivity in the shorter wavelengths. It requires pre-heating to make it conductive.

The globar is a low resistance rod of silicon carbide. It is useful in the longer wavelengths.

(*ii*) **The fore optics.** The beam is guided and focused by mirrors aluminized or silvered on their surfaces. Normally a focus is produced at the point where the sample is to be placed. Ordinary lenses and mirrors are not suitable as glass absorbs strongly over most of the frequencies used. Any windows which are essential (e.g., to contain a sample, or to protect the detector) must be made of mineral salts transparent to infra-red radiation which have been highly polished in order to reduce scattering to a minimum. NaCl, which is transparent above 650 $\text{cm}^{-1,}$ and KBr, transparent above 400 cm^{-1}, are much used. For aqueous samples, AgCl and CaF_2 sample cells can be used, and are infra-red transparent above 400 and 1200 cm^{-1}, respectively.

M_1 and M_2 divided the source radiation into two equivalent beams. One beam passes through the sample. The other beam passes through an equivalent path and is called the refernce beam. The two beams meet at the rotating sector mirror M_3. The rotating mirror alternately allows the sample beam through the spaces or reflects the refernce beam with a predetermined period to the monochromator slit. Consequently, the detector receives the sample beam and reference beam alternatly. Since both the beams have travelled the same path length, compensation for extraneous effects due to the atmosphere etc., have been taken care of.

(*iii*) **Monochromator.** It splits the polychromatic radiation into the component wavelengths by using prisms or gratings or both. Gratings give better resolving power. The prism or grating is rotatable mechanically. The resolution also depends on slit width and quality of mirrors.

(*iv*) **Detector and Amplifier.** Two main types of detectors are in common use, one sensing the heating effect of the radiation, the other depending on photoconductivity. In both the greater the effect (temperature or conductivity rise) at a given frequency, the greater the transmittance (and the less the absorbance) of the sample at that frequency.

An example of the temperature method is the Golay cell which is pneumatic in operation. The radiation falls on to a very small cell containing air, and temperature changes are measured in terms of pressure changes within the cell which can be recorded directly as 'transmittance'. Alternative examples of thermal detectors are small, sensitive thermocouples or bolometers.

Photoconductivity is the increase in electrical conductivity of a semiconductor caused by radiation incident on it. The radiation is allowed to fall on photoconducting material and the conductivity of the material is measured continously by a bridge network. The detector does not show any signal when the energy transmitted by reference and sample cells are equal. If the sample absorbs radiation there will be inequality in the beams and a signal is produced. The signal is then fed to a high gain amplifier which amplifies the low intensity signal upto a useable level with minimum thermal and electrical noise.

Photoconductive detectors in common use are indium antimonide (InSb) which can be used above 1400cm^{-1} and mercury cadmium telluride (MCT) used above ~ 700cm^{-1}. These detectors operate at liquid nitrogen temperatures, which reduces noise, and have a faster response time and sensitivity. Where extremely high sensitivity is required doped photoconductors such as copper-or gold-doped germanium detectors can be used-but these generally need to operate at liquid helium temperatures.

(*v*) **Recorder.** The amplified signal is used to move an attenuator (comb) which cuts down the radiation coming out of the reference beam until energy balance is restored. This is achieved by a motor which drives the comb into reference beam when an absorbing band is encountered and out of the beam when the band is passed over. The recorder pen is also coupled to this motor so that the comb movement is followed exactly by the pen.

Advantages. (*i*) By balancing sample and reference beams, the absorption of atmospheric CO_2 and H_2O does not appear in the infra-red spectra since both beams are reduced in energy to the same extent.

(*ii*) It is much simpler to amplify the alternating signal produced than the d.c. signal resulting from a single-beam detector.

(*iii*) The sector mirror acts as a modulator since it interrupts the beam periodically and by amplifying only that component of the signal having the sector mirror frequency (usually 10-100 rotations per second), a great improvement in the signal-to-noise ratio results.

(*iv*) When examining the spectra of solutions, one can put a cell containing the appropriate quantity of pure solvent into the reference beam, thus eliminating the solvent spectrum from the final trace. On a single-beam instrument the solvent spectrum must be taken separately and 'subtracted' from the solution spectrum in order to arrive at the spectrum of the substance of interest.

Disadvantages. (*i*) A double-beam instrument is never completely effective in removing traces of water vapour or CO_2 from the spectra. No matter how carefully the instrument is assembled small differences occur in the beam paths and a small residual spectrum results. This can usually be removed however by sweeping with dry, inert gas as well as using the double-beam principle.

(*ii*) The double-beam instrument only removes the spectral trace of CO_2 and H_2O. The very strong absorption of energy by these gases still remains in both beams. This means that at some parts of the spectrum the actual amount of energy reaching the detector may be extermely small. Under these conditions, unless the spectrometer is very carefully operated, the spectral trace of a substance may be quite false. Fortunately, regions of very high atmospheric absorption are few and narrow but they should be borne in mind when examinig infra-red spectra. This disadvantage can only be removed by sweeping out or evacuating the spectrometer. Similar, but more pronounced, effects occur in regions of strong solvent absorbance when a compensating cell is put in the reference beam.

24.16 Complementary Character of Raman and Infra-red Spectra

The differences between the Raman spectra and IR spectra are given below:

Raman Spectra	***IR Spectra***
1. It appears due to the scattering of radiation by the vibrating molecules.	1. It arises due to the absorption of radiation by the vibrating molecules.
2. The polarizability of the molecule changes.	2. The dipole moment of the molecules changes.
3. The intensity of the Raman line depends on the magnitude of the change of polarizability of the molecule.	3. The intensity of an IR absorption band depends on the magnitude of the change in dipole moment of the molecule.
4. Homopolar molecules respond to Raman scattering.	4. Homopolar molecules cannot exhibit IR spectra.

For IR spectra to occur, there must be a change in the dipole moment of the molecule. For the Raman spectra to occur, there must be a change in the polarizability of the molecule. Since these two requirements are somewhat different, lines may appear in one of the spectra or in both.

The symmetrical stretchings of the molecule which are usually missing in the IR spectra appear very prominently in Raman spectra, while the asymmetric vibrations exhibit opposite behaviour. It follows that the vibrational modes which are inactive in IR spectra, are sometimes active in Raman spectra. For example, in the case of CO_2 (a liner molecule),

$$\underset{\text{Sym. stretching}}{O \overset{\leftarrow}{=} C \overset{\rightarrow}{=} O} \begin{cases} \text{No change in dipole moment} & \boxed{\text{IR inactive}} \\ \text{Appreciable change in polarizability} & \boxed{\text{Raman active}} \end{cases}$$

$$\underset{\text{A sym. stretching}}{\vec{O} = C = \vec{O}} \begin{cases} \text{Appreciable change in dipole moment} & \boxed{\text{IR active}} \\ \text{No change in polarizability} & \boxed{\text{Raman inactive}} \end{cases}$$

$$\left.\begin{array}{c} \uparrow O = C = O \uparrow \\ \text{In-plane bending} \\ \overset{+}{O} = \overset{-}{C} = \overset{+}{O} \\ \text{Out-of-plane bending} \end{array}\right\} \text{Both are degenerate} \quad \begin{array}{c} \boxed{\text{IR active}} \\ \boxed{\text{Raman inactive}} \end{array}$$

With small molecules, the less symmetric vibrations often produce intense IR-bands, while the (more) symmetric vibrations often give rise to strong Raman bands.

According to *Rules of Mutual Exclusion*, which states that if a molecule has a centre of symmetry (e.g., centrosymmetric molecule like H_2, CO_2 , C_2H_2, etc.), the Raman active vibrations are IR-inactive and *vice versa*. The molecule for which the centre of symmetry is absent, may be simultaneously Raman and IR active. It is, therefore, the stretching vibrations of homonuclear diatomic molecules (such as H_2, N_2, O_2, etc.) which are IR-inactive, are Raman active.

Thus, it may be stated that the IR and Raman spectra are complementary to each other.

RAMAN SPECTROSCOPY

1. Raman Spectrometer

Fig. 24.12 shows the schematic diagram of a Raman spectrometer.

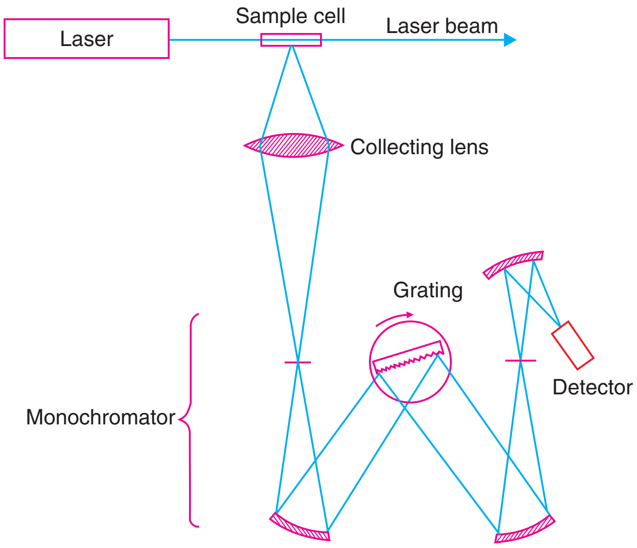


Fig. 24.12

The laser beam is passed through a cell, usually a narrow glass or quartz tube filled with the sample. Here the major considerations are the effective illumination of the sample and collection of scattered radiation from the sample. Light scattered sideways from the sample is collected by a lens and passed into a grating monochromator. The signal is measured by a sensitive photomultiplier and after amplification, it is processed by a computer which plots the Raman spectrum.

2. Classical Theory of Raman Effect

When a molecule is placed in an electric field E, the field distorts the electronic structure and polarizes the molecule. The polarization induced (P) is proportional to the applied electric field (E).

$$P = \alpha E \qquad \text{...(1)}$$

Here, α is the polarizability of the molecule.

When radiation of frequency ν_0 is allowed to fall on the molecules, each molecule experiences a varying electric field E given by

$$E = E_0 \cos 2\pi\nu_0 t \qquad \text{...(2)}$$

For a vibrating molecule, the polarizability along a direction will vary about an average value α_0 with a frequency equal to the vibrational frequency. We can write

$$\alpha = \alpha_0 + \beta \cos 2\pi\nu_m t \qquad \text{...(3)}$$

Here, α_0 is the equilibrium polarizability, and β is the rate of change of polarizability with the vibration. Then the induced polarization is given by

$$P = (\alpha_0 + \beta \cos 2\pi\nu_m t)(E_0 \cos 2\pi\nu_0 t) \qquad \text{...(4)}$$

$$\therefore\ P = \alpha_0 E_0 \cos 2\pi\nu_0 t + \frac{E_0\beta}{2}[\cos 2\pi(\nu_0 + \nu_m)t + \cos 2\pi(\nu_0 - \nu_m)t] \qquad \text{...(5)}$$

Thus, the induced polarization contains the three distinct frequency components ν_0, $(\nu_0 + \nu_m)$ and $(\nu_0 - \nu_m)$.

ν_0 is the frequency of incident radiation *i.e.*, Rayleigh scattering.

The last two frequencies are more and less than the frequency of incident radiation and predict the existence of Raman scattering.

No Raman line will be observed, if the change in polarizability β is zero.

Thus we have the general rule:

In order to be Raman active a molecular rotation or vibration must cause some change in a component of the molecular polarizability.

3. Vibrational Raman Spectra of Diatomic Molecules

The transition energies of the vibrational Raman spectra for a diatomic molecule can be written using the energy expression of an anharmonic oscillator [cf. Eq. (2), Section 23.14, Chapter 23].

$$\varepsilon = \bar{\nu}_e\left(V+\frac{1}{2}\right)-\bar{\nu}_e x_e\left(V+\frac{1}{2}\right)^2 m^{-1}, \text{ for } V = 0, 1, 2.... \qquad ...(1)$$

Here, $\bar{\nu}_e$ is called the *equilibrium oscillation frequency* of the molecule expressed in wave numbers and is the value for small displacements. x_e is the corresponding anharmonicity constant.

The selection rule for Raman scattering is given by $\Delta V = \pm 1, \pm 2$, ...

Considering the $V = 0 \rightarrow V = 1$ transition, we have

$$\Delta\varepsilon = \bar{\nu}_e(1-2x_e)\, m^{-1} \qquad ...(2)$$

During collision, if the molecule gains energy from the exciting radiation $\bar{\nu}_0$, it gives rise to a line on the low frequency side of $\bar{\nu}_0$ which is the Stokes Raman line.

In anti-Stokes Raman, the exciting radiation gains energy from the molecule. Hence, we may write

$$\bar{\nu} = \bar{\nu}_0 \pm \bar{\nu}_e(1-2x_e) \qquad ...(3)$$

Here the minus sign represents the Stokes' lines. The plus sign refers to the anti-Stoke's lines. Since the Raman scattering of light is of low intensity, we can ignore completely all the weaker effects such as overtones and '*hot*' bands.

4. Pure Rotational Raman Spectra (Linear Molecules)

The rotational energy levels of linear molecules are given by [cf. Eq. (6), Section 23.9, Chapter 23]

$$\varepsilon_J = BJ(J+1) - DJ^2(J+1)^2 \text{ cm}^{-1} \ (J = 0, 1, 2,....)$$

But, in Raman spectroscopy, the precision of the measurements does not normally warrant the retention of the term involving D, the centrifugal distortion constant. Thus we take the simpler expression:

$$\varepsilon_J = BJ(J+1) \text{ cm}^{-1} \ (J = 0, 1, 2,...) \qquad ...(1)$$

to represent the energy levels.

Transitions between these levels follow the selection rule:

$$\Delta J = 0, \text{ or } \pm 2 \text{ only} \qquad ...(2)$$

We define ΔJ as $(J_{\text{upper state}} - J_{\text{lower state}})$. We can ignore the selection rule $\Delta J = -2$ since, for a pure rotational change, the upper state quantum number must necessarily be greater than that in the lower state. Further, the 'transition' $\Delta J = 0$ is trivial since this represents no change in the molecular energy and hence Rayleigh scattering only.

Combining, then, $\Delta J = +2$ with the energy levels of Eq. (1) we have:

$$\begin{aligned}\Delta\varepsilon &= \varepsilon_{J' = J+2} - \varepsilon_{J'' = J} \\ &= B(4J + 6)\ \text{cm}^{-1}\ (J = 0, 1, 2, \ldots) \qquad \ldots(3)\end{aligned}$$

The wavenumbers of the rotational Raman lines are given by

$$\bar{\nu} = \bar{\nu}_0 \pm B(4J + 6)\ \text{cm}^{-1} \qquad \ldots(4)$$

Here, the plus sign refers to anti-Stokes' lines, the minus to Stokes' lines, and $\bar{\nu}_0$ is the wavenumber of the exciting radiation.

Stokes' and anti-Stokes' lines have comparable intensity because many rotational levels are populated and hence downward transitions are approximately as likely as upward ones.

Fig. 24.13 shows the rotational energy levels of a diatomic molecule and the rotational Raman spectrum arising from transitions between them. Spectral lines are numbered according to their lower J values.

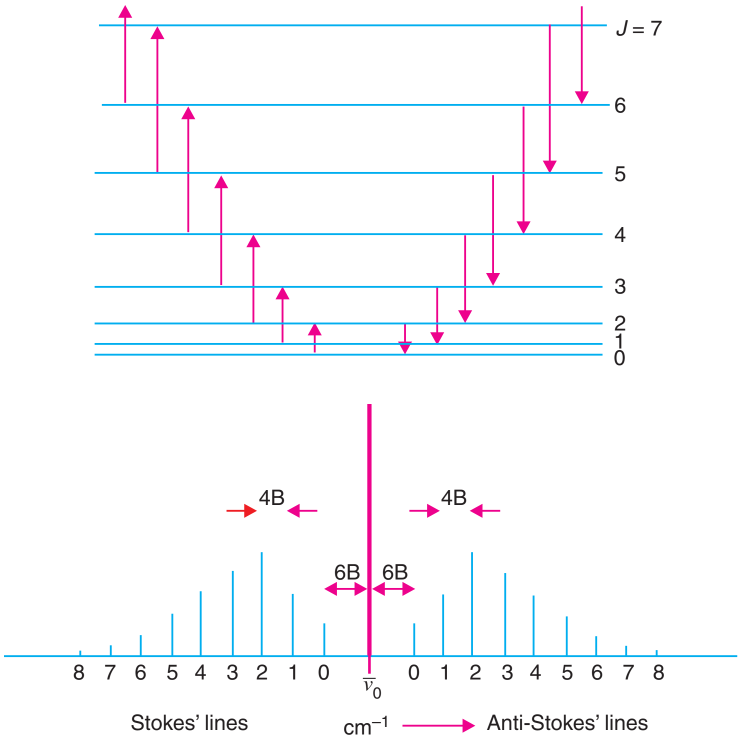


Fig. 24.13

Putting $J = 0$ in Eq. (4), we find that the separation of the first line from the exciting line is $6B$ cm^{-1}, while the separation between successive lines is $4B$ cm^{-1}.

For diatomic and light triatomic molecules the rotational Raman spectrum will normally be resolved and we can obtain a value of B. Hence the moment of inertia and bond length for such molecules can be found.

Homonuclear diatomic molecules (for example O_2, H_2) give no infra-red or microwave spectra since they possess no dipole moment, whereas they *do* give a rotational Raman spectrum. Therefore, structure of such molecules can be obtained from Raman technique. Raman technique is thus complementary to microwave and infra-red studies.

If a molecule has a centre of symmetry (*e.g.*, H_2, O_2, CO_2), then the effects of nuclear spin will be observed in both the Raman and infra-red spectra. Thus for O_2 and CO_2 (since the spin of oxygen is zero) every alternate rotational level is absent.

For example, in the case of O_2, every level with *even J* values is missing, and thus every transition labelled $J = 0, 2, 4, \ldots$ in Fig. 24.13 is also completely missing from the spectrum. In the case of H_2, and other molecules composed of nuclei with non-zero spin, the spectral lines show an alternation of intensity.

EXERCISE

1. Describe and explain the working of a "Constant Deviation Spectrograph".

2. Describe any suitable apparatus for the ultraviolet spectroscopy. Discuss the peculiarities of the different components of the apparatus used for the purpose. To what wavelength could this be used? How would you go beyond ? (*Calcutta, Hons.*)

3. Give a brief account of any method of producing and studying the infra-red spectrum. What is the importance of such a study ? (*Burdwan, Hons.*)

4. The first rotational Raman line of H_2 appears at 34600 m^{-1} from the exciting line. Calculate the bond length of H_2 molecule, Given $^1H = 1.673 \times 10^{-27}$ kg.

Solution : The first rotational Raman line appears at a distance of 6B from the exciting line.

$$6\text{ B} = 34600\text{ m}^{-1};\ \text{B} = 5766.7\text{ m}^{-1} = \frac{}{Ic}$$

$$I = \frac{h}{8\pi^2 cB} = \frac{6.626\times10^{-34}\text{ Js}}{8\pi^2(3\times10^8\text{ m/s})(5766.7\text{m}^{-1})} = 4.8514\times10^{-48}\text{kg m}^2$$

$$\text{Reduced mass, } \mu = \frac{m_1 m_2}{m_1+m_2} = \frac{(1.673\times10^{-27}\text{ kg})(1.673\times10^{-27}\text{ kg})}{2\times1.673\times10^{-27}\text{ kg}} = 0.8365\times10^{-27}\text{ kg}$$

$$r^2 = \frac{I}{\mu} = \frac{4.8514\times10^{-48}\text{kg m}^2}{0.8365\times10^{-27}\text{kg}} = 0.5800\times10^{-20}\text{m}^2$$

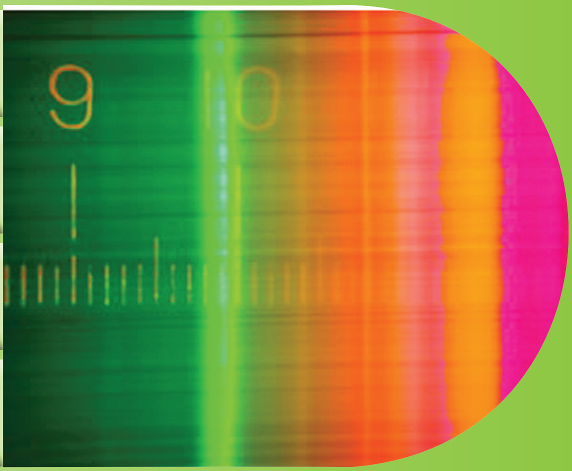
$\therefore\ r = 0.762 \times 10^{-10}$ m

5. The pure rotational levels of a molecule in the far-infrared region follows the formula F(J) = BJ (J + 1), where F(J) is the energy of the rotational level with quantum number J and B is the rotational constant. The lowest rotational energy gap in rotational Raman spectrum is

(*a*) 2B (*b*) 4B (*c*) 6B (*d*) 8B **Ans.** (*b*)

CHAPTER

25



SPECTRA OF ALKALI ELEMENTS

AT A GLANCE

25.1 Introduction
25.2 Different Series in Alkali Spectra: Main Features
25.3 Energy Levels in Alkali Spectra and Quantum Defect
25.4 Nonpenetrating and Penetrating Orbits
25.5 Explanation of Salient Features of Alkali Spectra
25.6 The Spectrum of Lithium and Other Hydrogen like Species
25.7 The Spectrum of Helium and the Alkaline Earths

25.1 Introduction

The main examples of alkali metals are lithium, sodium, potassium, rubidium and cesium. In case of alkali metals, the outermost shell contains a single electron. This electron is called the 'optical' or 'valence' electron. The spectral lines of an alkali atom arise due to the transitions of this single valence electron. The alkali spectrum is called 'one-electron' spectrum.

25.2 Different Series in Alkali Spectra: Main Features

The alkali spectrum consists of spectral lines which can be classified into four series : principal series, sharp series, diffuse series and fundamental series. The principal series is the most prominent and can be observed in emission as well as in absorption spectrum. The other series are observed in emission spectrum only.

1. Principal series. The series arises from the transitions between various *P*-levels and lowest *S*-level. The lowest *S*-level has the lowest possible value for the energy and represents the ground state of the atom. The wave number of the series is given by the relation.

$$\bar{\nu}_p = \frac{R}{(1+\mu_s)^2} - \frac{R}{(m+\mu_p)^2}, \quad m \geq 2 .$$

2. Sharp series. This type of series comes from the transitions from *S*-levels (exclusive of the lowest) to the lowest of the *P*-levels.

The wave number of the series is given by the relation

$$\bar{\nu}_s = \frac{R}{(2+\mu_p)^2} - \frac{R}{(m+\mu_s)^2}, \quad m \geq 2 .$$

3. The diffuse series. It arises from the transitions between the various *D*-levels and the lowest *P*-level.

The wave number of the series is given by the relation

$$\bar{\nu}_D = \frac{R}{(2+\mu_p)^2} - \frac{R}{(m+\mu_D)^2}, \quad m \geq 3 .$$

4. Fundamental series or bergmann series. The fundamental series arises from the transitions from various *F*-levels to the lowest *D*-level.

The wave number of the series is expressed by the relation

$$\bar{\nu}_F = \frac{R}{(3+\mu_D)^2} - \frac{R}{(m+\mu_F)^2}, \quad m \geq 4 .$$

Here μ_S, μ_P, μ_D and μ_F are the characteristic constants of sharp, principal, diffuse and fundamental series, respectively. The abbreviated forms for writing the wave number of each series are as follows:

1. $(1, \mu_S) \leftarrow (m, \mu_p)$ with $m \geq 2$	Principal series.
2. $(2, \mu_S) \leftarrow (m, \mu_s)$ with $m \geq 2$	Sharp series.
3. $(2, \mu_p) \leftarrow (m, \mu_D)$ with $m \geq 3$	Diffuse series.
4. $(3, \mu_D) \leftarrow (m, \mu_F)$ with $m \geq 4$	Fundamental series.

The other important features of the series are:

1. The sharp and diffuse series have a common limit.
2. The wave number interval between this common limit and the limit of the principal series is equal to that of first line of principal series.

25.3 Energy Levels in Alkali Spectra and Quantum Defect

The emission of alkali spectral lines can be fairly explained on the same lines as the Bohr-Sommerfeld theory for hydrogen atom. An atom has a number of discrete energy states. Each state is characterised by a total quantum number n (= 1, 2, 3, ... ∞). For each value of n, there are component levels labelled by an additional quantum number l, called the 'orbital' quantum number. l can take values, 0, 1, 2, ... $(n-1)$. Thus $n = 1$ state has only one level ($l = 0$); $n = 2$ state has two levels ($l = 0, 1$), and so on. The levels corresponding to $l = 0, 1, 2, 3$, ... are called *s*, *p*, *d*, *f*, ... levels respectively. Thus $n = 1$ state has a level called 1*s*; the $n = 2$ state has two levels called 2*s* and 2*p*; the $n = 3$ state has three levels called 3*s*, 3*p* and 3*d*; the $n = 4$ state has four levels 4*s*, 4*p*, 4*d*, 4*f*, and so on.

The energies of these levels are given by

$$E_{n,l} = -\frac{Rhc}{(n-\Delta)^2}.$$

Here, Δ is called 'quantum defect' and depends on l. Thus the energies of levels with same n but different l are *different*.

25.4 Nonpenetrating and Penetrating Orbits

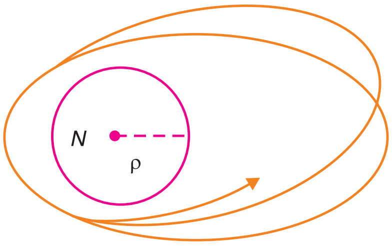


Fig. 25.1

***(i)* Nonpenetrating Orbits.** We assume that the charge due to the electrons in the closed core is distributed uniformly over a hollow sphere of radius ρ with the nucleus (*N*) at its centre (Fig. 25.1). Suppose in the course of its motion round the nucleus, the elliptic orbit of the single optical electron in an alkali atom of atomic number *Z* does not penetrate the closed orbits or core of $(Z-1)$ electrons. We expect the energy of the optical electron in the orbit to be nearly the same as that of the electron in the corresponding orbit in a hydrogen atom, on the Bohr-Sommerfeld model. It is found that the energies of the *F* states of the sodium atom agree to within one percent with the corresponding energy values of the hydrogen atom. Such orbits for which the energies are nearly the same as in hydrogen atoms are called *nonpenetrating orbits*.

Nonpenetrating orbits are defined as those orbits for which the observed energies are very nearly equal to those of the corresponding hydrogen-like orbits.

The f orbits in all of the alkali metals are good examples of nonpenetrating orbits.

The resultant field under which the optical electron describes the orbit varies slowly as this electron moves in its orbit about the nucleus and the closed electron core. So the elliptic Bohr-Sommerfeld orbits in these cases precess slowly round the nucleus.

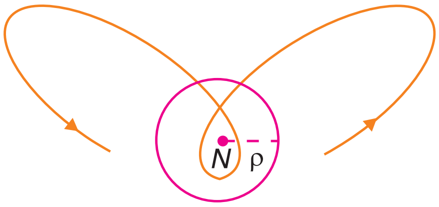


Fig. 25.2

***(ii)* Penetrating Orbits.** If the optical electron orbit in an alkali atom penetrates the completed electron shells or sub-shells, more and more of the closed electron core is left behind as the optical electron moves in its orbit from aphelion to perihelion. This results in an increase in the field under which this electron traces its path as its distance from the nucleus goes from the maximum to the minimum value (Fig. 25.2). So the electron traverses a path of increasing eccentricity. The increase in the force of attraction on the optical electron makes its energy in the orbit less than that for a corresponding Bohr-Sommerfeld state in a hydrogen atom. Thus the energies of the *S* and *P* states in the sodium atom are less than those of the corresponding states of the hydrogen atom. Such orbits for which the energies are considerably different from the corresponding values in the hydrogen atom are called *penetrating orbits*.

25.5 Explanation of Salient Features of Alkali Spectra

The electronic configuration of alkali metals is such that the core of an inert gas is surrounded by an *s* electron.

$$\begin{aligned}\text{Li} &= [\text{He}] + 2s^1\\ \text{Na} &= [\text{Ne}] + 3s^1\\ \text{K} &= [\text{Ar}] + 4s^1\\ \text{Rb} &= [\text{Kr}] + 5s^1\\ \text{Cs} &= [\text{Xe}] + 6s^1\end{aligned}$$

The core electrons do not play any part in optical spectra. Only valence electrons, also called *optical electron* determine the characteristics of optical spectra. The configuration is thus identical to hydrogen.

Let us discuss the characteristics of alkali spectra taking sodium as a typical example of an alkali element. The optical 3*s* electron of sodium when excited jumps to higher energy states, such as 3*P*, 4*S*, 3*D*, 4S, 5*S*, 4*D* ... etc. From the higher level the electron jumps back to the lower levels but only such that the *l* value changes by ±1 *i.e.,*

$$\Delta l = \pm 1.$$

This is called the '*selection rule*'.

Fig. 25.3 gives the energy level diagram of the sodium atom and shows the allowed transitions of the optical electron.

(*i*) When the electron jumps from any *p*-level to the lowest *s*-level (3*s*), it emits a line of principal series.

(*ii*) When the electron jumps from any *s*-level to the lowest *p*-level (3*p*), it emits a line of sharp series.

(*iii*) When the electron jumps from any *d*-level to the lowest *p*-level (3*p*), it emits a line of diffuse series.

(*iv*) When the electron jumps from any *f*-level to the lowest *d*-level (3*d*), it emits a line of fundamental series.

Thus emission of all the spectral series is explained.

The different series in the spectrum arise from the following transitions of the optical electron.

Principal series	$n\ {}^2P \rightarrow 3^2S,\quad n = 3, 4, 5...$
Sharp series	$n\ {}^2S \rightarrow 3^2P,\quad n = 4, 5, 6\ ...$
Diffuse series	$n\ {}^2D \rightarrow 3^2P,\quad n = 3, 4, 5\ ...$
Fundamental series	$n\ {}^2F \rightarrow 3^2D,\quad n = 4, 5, 6\ ...$

For a given set of values of the principal quantum number, we get several series here, unlike in the case of hydrogen.

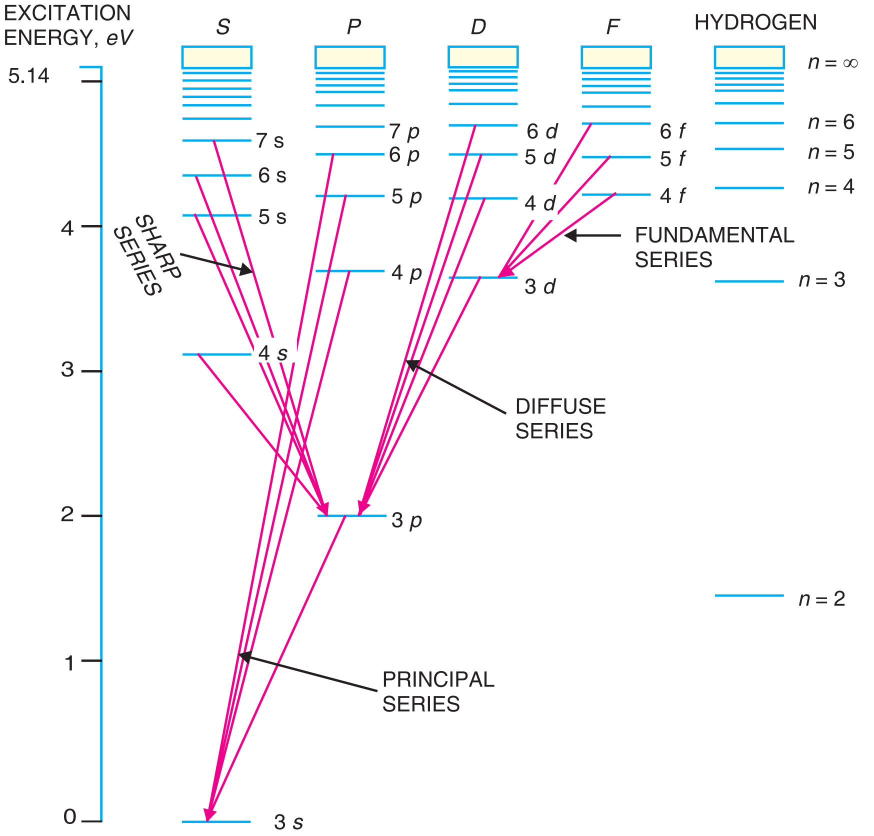

EXCITATION ENERGY, *eV*
5.14
4
3
2
1
0
S
P
D
F
HYDROGEN
n = ∞
7 s
6 s
5 s
4 *s*
3 *s*
7 *p*
6 *p*
5 *p*
4 *p*
3 *p*
6 *d*
5 *d*
4 *d*
3 *d*
6 *f*
5 *f*
4 *f*
n = 6
n = 5
n = 4
n = 3
n = 2
SHARP SERIES
FUNDAMENTAL SERIES
DIFFUSE SERIES
PRINCIPAL SERIES

Fig. 25.3

According to Bohr-Sommerfeld theory and quantum mechanical theory, all substates belonging to a given n in hydrogen have same value of energy. But this l degeneracy is removed in a multi-electron system because of shield of nuclear charge and penetration of atomic core. The term value is written as

$$T = \frac{RZ_0^2}{n_{eff}^2}$$

Here, Z_0 is effective nuclear charge outside the core. $n^2{}_{\text{eff}}$ is effective quantum number. As Z_0 is greater than one and $n_{\text{eff}} < n$, the term value is increased. As a result, energy levels in alkali atoms lie lower than the corresponding hydrogen levels. However, for large n, Z_0 is equal to unity and $n_{\text{eff}} \rightarrow n$. The corresponding levels approach the hydrogen levels.

25.6 The Spectrum of Lithium and Other Hydrogen-like Species

The energy levels of the lithium atom are sketched in Fig. 25.4. It shows the difference in energy of s, p, and d states with the same value of n. Some allowed transitions are also shown.

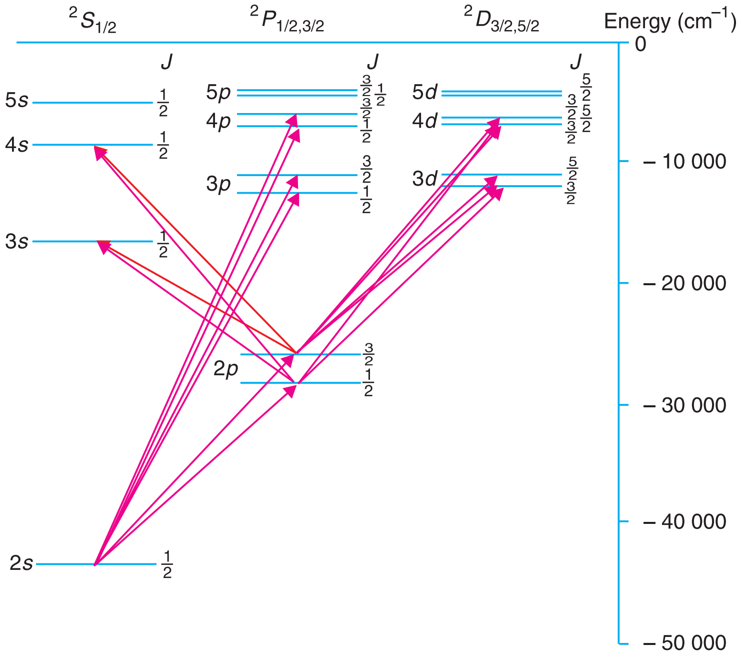


Fig. 25.4

Selection Rules : The selection rules for alkali metals are the same as for hydrogen.

$$\Delta n = \text{anything}, \ \Delta l = \pm 1, \ \Delta j = 0, \pm 1$$

(*i*) The transitions from the ground state ($1s^2\ 2s$) can occur to p levels:

$$2S_{1/2} \longrightarrow nP_{1/2,\,3/2}$$

A series of doublets similar to the Lyman series will be formed.

(*ii*) From the $2p$ state, two separate series of lines will be seen:

$$2^2P_{1/2,\,3/2} \longrightarrow n^2S_{1/2}$$

and $$2^2P_{1/2,\,3/2} \longrightarrow n^2D_{3/2,\,5/2}$$

The former will be doublets, the latter compound doublets, but their frequencies will differ because the *s* and *d* orbital energies are no longer the same.

The same remarks apply to the other alkali metals, the differences between their spectra and that of lithium being a matter of scale only. For instance the *j*-splitting due to coupling between *l* and *s* increases markedly with the atomic number. Thus the doublet separation of lines in the spectral series, which is scarcely observable for hydrogen, is less than 1cm^{-1} for the 2*p* level of lithium, about 17 cm^{-1} for sodium, and over 5000 cm^{-1} for cesium.

Any atom which has a single electron moving outside a closed shell will exhibit a spectrum of the type discussed above. Thus ions of the type He$^+$, Be$^+$, B^{2+}, etc., show 'hydrogen-like spectra'.

25.7 The Spectrum of Helium and the Alkaline Earths

Helium consists of a central nucleus and two outer electrons. There are only two possibilities for the relative spins of the two electrons:

Case 1. Singlet states

The spins of the two electrons are paired (↑↓). If S_{1z} is $+\dfrac{1}{2}$, S_{2z} must be $-\dfrac{1}{2}$.

Hence $S_z = S_{1z} + S_{2z} = 0$. So S = 0 and we have *singlet* states.

The lowest possible energy state of this atom is when both electrons occupy the 1*s* orbital. By Pauli's principle, this is possible only if their spins are paired. So the ground state of helium must be a singlet state. Further, L = $l_1 + l_2 = 0$, and hence *J* can only be zero. The ground state of helium, therefore, is $^1S_0$.

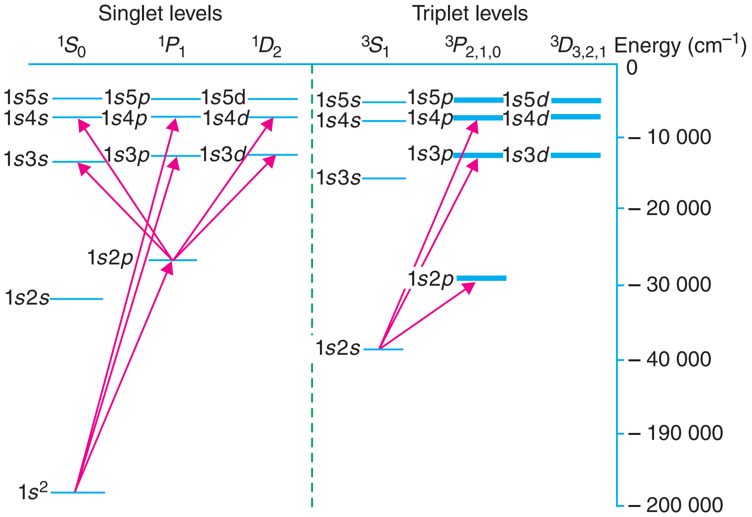


Fig. 25.5

The selection rules for many-electron systems are:

$$\Delta S = 0 \quad \Delta L = \pm 1 \quad \Delta J = 0, \pm 1 \qquad \text{...(1)}$$

Since *S* cannot change during a transition, the singlet ground state can undergo transitions only to other singlet states.

We assume that only one electron undergoes transitions, leaving the other in the 1*s* orbital.

The left-hand side of Fig. 25.5 shows the energy levels of the electrons and the allowed transitions.

(*i*) Initially the $1s^2\ {}^1S_0$ state can undergo a transition only to $1s^1\ np^1$ states (abbreviated to $1snp$). In the $1snp$ state $L = 1$, $S = 0$, and hence $J = 1$ only.

So the transition may be symbolized :

$$1s^2\ {}^1S_0 \longrightarrow 1snp\ {}^1P_1$$

or, briefly: $${}^1S_0 \longrightarrow {}^1P_1$$

(*ii*) From the ${}^1P_1$ state the system could either revert to ${}^1S_0$ states, or undergo transitions to the higher ${}^1D_2$ states (for these $S = 0$, $L = 2$ and hence $J = 2$ only).

In general, then, all these transitions will give rise to spectral series very similar to those of lithium except that here transitions are between *singlet* states only and all the spectral lines will be single.

Case 2. Triplet states

The electron spins are parallel (↑↑). Now $S_{1z} = S_{2z} = +\dfrac{1}{2}$. So $S = 1$ and the states are *triplet.*

Since the electrons are now forbidden by Pauli's principle from occupying the same orbital, the lowest energy state is $1s\ 2s$. This and other triplet energy levels are shown on the right of Fig. 25.5. The $1s\ 2s$ state has $S = 1$, $L = 0$, and hence $J = 1$ only, and so it is ${}^3S_1$. By the selection rules of Eq. (1) it can only undergo transitions into the $1snp$ triplet states. These, with $S = 1$, $L = 1$, have $J = 2$, 1, or 0. So, the transitions may be written:

$${}^3S_1 \rightarrow {}^3P_2,\ {}^3S_1 \rightarrow {}^3P_1 \text{ and } {}^3S_1 \rightarrow {}^3P_0$$

All three transitions are allowed, since $\Delta J = 0$ or ± 1, so the resulting spectral lines will be *triplets.*

Transitions between ^{3}P and ^{3}D Levels

Transitions from the ${}^3P$ states may take place to ${}^3D$ states. For ${}^3D$ we have $S = 1$, $L = 2$, and hence $J = 3$, 2, or 1. Fig. 25.6, shows the 'compound triplet', spectrum arising from transitions between ${}^3P$ and ${}^3D$ states, bearing in mind the selection rule $\Delta J = 0, \pm 1$.

(*i*) ${}^3P_2$ can go to each of ${}^3D_{3,\,2,\,1}$,

(*ii*) ${}^3P_1$ can go only to ${}^3D_{2,\,1}$, and

(*iii*) ${}^3P_0$ can go only to ${}^3D_1$

Thus the complete spectrum should consist of six lines. Normally, however, the very close spacing is not resolved, and only three lines are seen. For this reason the spectrum is referred to as a *compound triplet*.

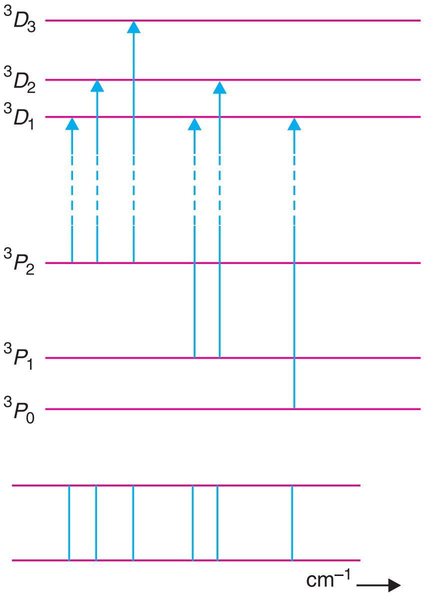


Fig. 25.6

We see, then, that the spectrum of helium consists of spectral series grouped into two types which overlap each other in frequency.

(*i*) In one type, involving transitions between singlet levels, all the spectral lines are themselves singlets.

(*ii*) In the other type, the transitions are between triplet states and each 'line' is at least a close triplet and possibly even more complex.

Because of the selection rule $\Delta S = 0$ there is a strong prohibition on transitions between singlet and triplet states, and transitions *cannot occur* between the right- and left-hand sides of Fig. 25.5.

Other atoms containing two outer electrons exhibit spectra similar to that of helium. Thus the alkaline earths, beryllium, magnesium, calcium, etc., fall into this category, as do ionized species with just two remaining electrons, for example, B^+ , C^{2+}, etc.

EXERCISE

1. Describe the general features of the spectra of alkali-like atoms. How are they explained? In what respect do they differ from the spectra of H_2 ? (*Purvanchal University*, 90, 91)
2. What are penetrating and non-penetrating orbits ? (*P.U.* 91)

CHAPTER

26

X-RAY SPECTRA

AT A GLANCE	
26.1	Fine Structure of X-ray Levels
26.2	Spin-relativity Doublets (Regular Doublets)
26.3	The Regular-doublet Law
26.4	Screening Doublets
26.5	Irregular-doublet Law
26.6	X-ray Absorption Spectrum
26.7	Auger Effect

26.1 Fine Structure of X-ray Levels

The resolution of L level into 3 sublevels, of M level into 5 sublevels and of N level into 7 sub-levels is called *fine structure* of X-ray levels. To account for the fine structure of X-ray levels, we consider the effect of electrons which screen the nuclear charge and the effect of spin of electrons.

The L electrons move in an approximately Coulomb field. The two K electrons act very much as if they had coalesced with the nucleus. This screening of the nucleus by the K electrons reduces the nuclear charge by 2.

This reducing of nuclear charge by the electrons which are nearer to the nucleus is called screening effect. In a similar way, for an electron in M shell, K and L electrons screen the nuclear charge.

Therefore, the energy of the system is written as

$$T = \frac{R(Z-b)^2}{n^2}.$$

Here, b is screening constant.

$(Z-b)$ may be called the effective atomic number of the element.

It is to be noted that the constant b also includes the effect of outer electrons on the inner ones. This effect is known as external screening. Further, when an electron moves in an orbit of low quantum number in the field of the nucleus of high atomic number, the velocity of the electron is high enough to take into account the relativistic variation of mass. With this argument, Sommerfeld and Wentzel showed that the energy of the system is calculated by

$$\begin{aligned} T &= \frac{R(Z-b)^2}{n^2} + \frac{R\alpha^2}{n^4}(Z-d)^4\left(\frac{n}{k}-\frac{3}{4}\right) \\ &\quad + \frac{R\alpha^4}{n^6}(Z-d)^6\left(\frac{1}{4}\frac{n^3}{k^3}+\frac{3n^2}{4k^2}-\frac{3n}{2k}+\frac{5}{8}\right) \\ &\quad + \frac{R\alpha^6}{n^8}(Z-d)^8\left\{\frac{1}{8}\frac{n^5}{k^5}+\frac{3}{8}\frac{n^4}{k^4}+\frac{1}{8}\frac{n^3}{k^3}-\frac{15}{8}\left(\frac{n}{k}\right)^2+\frac{15}{8}\frac{n}{k}-\frac{35}{64}\right\}. \end{aligned}$$

Here, $\alpha^2 = 4\pi^2 e^4/h^2c^2 = 5.30 \times 10^{-5}$ is the square of Sommerfeld's fine-structure constant, R is the Rydberg constant, b and d are screening constants, n is the total quantum number, and k is Sommerfeld's original azimuthal quantum number; $k = 1, 2, 3, \ldots$ for $s, p, d, \ldots$ electrons, respectively.

The first screening constant b depends on the distribution of electrons in shells both outside and inside the one containing the electron considered.

The second screening constant d is due to spin-orbit interaction and relativity.

The effect of two screening constants b and d in shifting the energy level from those of the hydrogen-like atom is illustrated in Fig. 26.1.

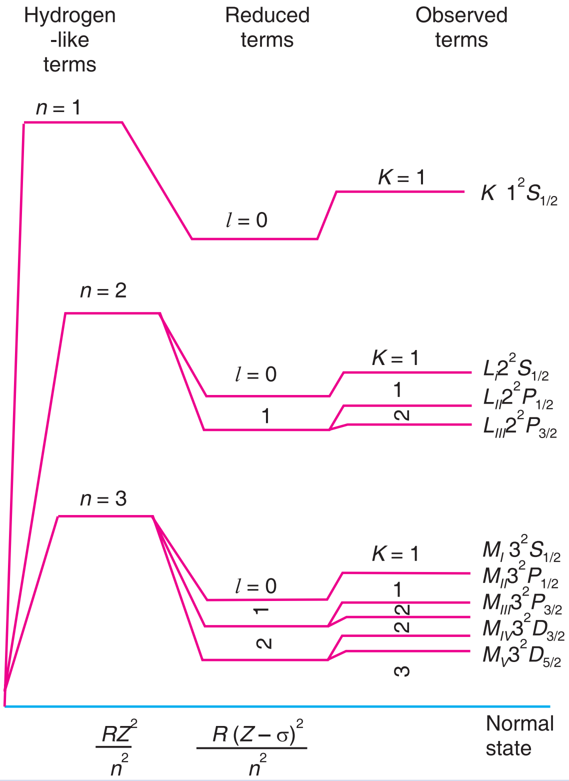

Hydrogen -like terms
Reduced terms
Observed terms
n = 1
l = 0
K = 1
K 1²S1/2
n = 2
l = 0
1
K = 1
1
2
L_I 2²S1/2
L_II 2²P1/2
L_III 2²P3/2
n = 3
l = 0
1
2
K = 1
1
2
2
3
M_I 3²S1/2
M_II 3²P1/2
M_III 3²P3/2
M_IV 3²D3/2
M_V 3²D5/2
RZ²/n²
R(Z − σ)²/n²
Normal state

Fig. 26.1

The effect of screening constant b is to break one level into n sub-levels and the corresponding term is known as reduced term. If now to this effect, we add the effect of second screening constant d, each sub-level again breaks up into two with the exception of ground level.

26.2 Spin-relativity Doublets (Regular Doublets)

The term *spin-relativity doublet* refers in general to each pair of energy levels having the same n, S, and L values but different J values. In x-ray notation these are the intervals $L_{II} - L_{III}$, $M_{II} - M_{III}$, $M_{IV} - M_V$, $N_{II} - N_{III}$, $N_{IV} - N_V$, $N_{VI} - N_{VII}$, $O_{II} - O_{III}$, $O_{IV} - O_V$, etc.

The term values for L_{II} and L_{III} are

$$\begin{aligned} T_{L_{II}} &= \frac{R(Z-d)^2}{2^2} + \frac{R\alpha^2}{2^4}(Z-d)^4 \\ &\quad \left\{\frac{5}{4} + \frac{21}{8}\frac{\alpha^2}{2^2}(Z-d)^2 + \frac{429}{64}\frac{\alpha^4}{2^4}(Z-d)^4 + \ldots\right\} \\ T_{L_{III}} &= \frac{R(Z-d)^2}{2^2} + \frac{R\alpha^2(Z-d)^4}{2^4} \\ &\quad \left\{\frac{1}{4} + \frac{1}{8}\frac{\alpha^2}{2^2}(Z-d)^2 + \frac{5}{64}\frac{\alpha^4}{2^4}(Z-d)^4 + \ldots\right\}. \\ \Delta\nu = T_{L_{II}} - T_{L_{III}} &= \frac{R\alpha^2}{2^4}(Z-d)^4 \left\{1 + \frac{5\alpha^2}{8}(Z-d)^2 + \frac{53\alpha^4}{128}(Z-d)^4 + \ldots\right\}. \end{aligned}$$

The value of d is constant for a particular value of L and it is independent of Z. For example, for $L_{II} - L_{III}$, the value of d is 3.50.

Neglecting higher terms,

$$\Delta\nu = \frac{R\alpha^2(Z-d)^4}{2^4}.$$

26.3 The Regular-doublet Law

The regular-doublet law refers to the law governing the spin-relativity doublets $L_{II} - L_{III}$, $M_{II} - M_{III}$, $M_{IV} - M_V$, $N_{II} - N_{III}$, etc.

Statement. *The doublet separation, $\Delta\nu$ in wave numbers, for any spin-relativity doublet is approximately proportional to the fourth power of the effective atomic number, $(Z - d)$.*

For $L_{II} - L_{III}$, we have,

$$\Delta\nu = \frac{R\alpha^2}{2^4}(Z-3.5)^4.$$

26.4 Screening Doublets

The doublet intervals $(L_I - L_{II})$, $(M_I - M_{II})$, $(M_{III} - M_{IV})$, $(N_I - N_{II})$, $(N_{III} - N_{IV})$... etc., are called screening doublets. These are the levels having same n, S, and J values but different L values.

Taking only the effect of screening into account, we write

$$T = \frac{R(Z-b)^2}{n^2}$$

or

$$\left(\frac{T}{R}\right)^{1/2} = \frac{Z-b}{n}.$$

Hence, we write the value $\left(\frac{T}{R}\right)^{1/2}$ for L_{I} and L_{II} as

$$\left(\frac{T}{R}\right)^{1/2}_{L_1} = \frac{Z-b_1}{n}$$

$$\left(\frac{T}{R}\right)^{1/2}_{L_n} = \frac{Z-b_2}{n}$$

$$\therefore \qquad \Delta\sqrt{\left(\frac{T}{R}\right)} = \frac{b_2-b_1}{n} = \frac{\Delta b}{n} = \text{constant.}$$

26.5 Irregular-doublet Law

Statement. *The difference between the square roots of the term values of a given doublet is a constant independent of the atomic number Z.*

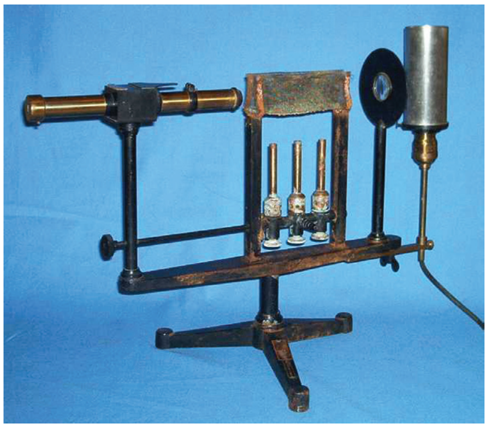
26.6 X-ray Absorption Spectrum

A beam of continuous X-rays is passed through the film of a substance. The transmitted beam is examined by an X-ray spectrograph. Then, the X-ray absorption spectrum of that substance is obtained. The nature of the X-ray absorption spectrum is quite different from the X-ray emission spectrum of the same substance.

(*i*) The emission lines K_α, K_β, K_γ….. do not appear in the absorption spectrum.

(*ii*) The absorption spectrum shows a continuous region bounded by a sharp edge (known as absorption edge) in the position of the limit of the *K*-series.

(*iii*) The absorption edges are characteristic of the absorber element. When successive elements in the periodic table are used as absorbers, the edge shifts toward higher frequency (Fig. 26.2). Thus these edges are characteristic of the absorber element and are called 'absorption edges'.



Absorption Spectrum.

(*i*) Explanation of Absence of Absorption lines. The K_α, K_β... emission lines cannot be observed as absorption lines. For example, K_α line in emission occurs when an electron drops from the *L*-shell to the *K*-shell. The line can occur in absorption if electron from *K*-shell may jump to *L*-shell after absorbing necessary energy from continuous beam. But this does not happen because the *L*-shell has already its full quota of 8 electrons and so it cannot receive any further electron (Pauli's exclusion principle). Hence the K_α-line cannot be observed as absorption line. This is true for all the *X*-ray emission lines that are commonly observed.

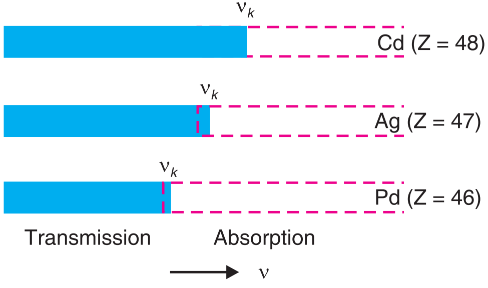


Fig. 26.2

(*ii*) Explanation of Appearance of Absorption Edges. An X-ray photon, capable of removing an inner electron to infinity (where no shells exist), can be absorbed. Let E_K be the energy required to remove a K-electron. Then, a photon of frequency $\nu \geq \nu_K$ can eject a K-electron where

$$h\nu_K = E_K.$$

Photons of frequency less than ν_K are transmitted by the absorbing element.

At the frequency ν_k, absorption in K-shell suddenly starts and continues for all frequencies greater than ν_K. Therefore, absorption spectrum is abruptly discontinued at ν_K and photographic plate is not blackened beyond ν_K.

As Z increases, E_K also increases. Hence ν_K also increases.

Multiplicity of Atomic Shells. When the X-ray absorption spectrum is taken in the L-region, again a continuum is observed but now with three absorption edges. This means that the L-shell has three energy levels from which electrons can be ejected to infinity. Similarly, M-shell has five energy levels, and so on. Thus X-ray absorption spectrum reveals the multiplicity of atomic shells.

26.7 Auger Effect

We have seen that when a K-electron is knocked out from an atom either by high-energy electron impact or through absorption of an X-ray photon, the vacancy thus created in the K-shell is filled up by transition from an outer-shell (say, L) electron. The energy of the emitted characteristic x-ray photon is $h\nu_{KL} = E_K - E_L$.

But Auger found that there are certain radiationless transitions.

The excess energy in such transitions may be directly absorbed by another L-electron which gets emitted. There are now two vacancies created in L-shell and the energy level of the atom is different from L-level. If we designate this as LL-level of energy E_{LL}, then the kinetic energy of the emitted second electron is given by

$$\frac{1}{2}mv^2 = E_K - E_{LL}.$$

Such a transition in which there is no emission of electromagnetic radiation, but emission of two electrons from the same atom is called **Auger transition or Auger effect.** It is also called *radiationless transition*.

Thus, the de-excitation of the atom may be accompanied either by the emission of a photon (characteristic radiation) or an electron (Auger electron).

EXERCISE

1. What are spin and screening doublets in X-ray spectra? Explain their origin.

   or

   Discuss the occurrence of doublets in X-ray spectra.

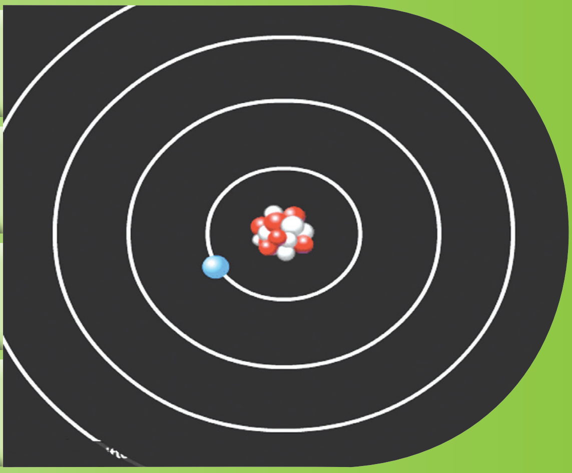
   or

   Analyse the fine structure of X-ray energy levels and give an account of screening and spin relativity doublets.
2. Discuss X-ray absorption spectrum and the importance of sharp absorption edges.
   (*Gorakhpur* 87, *Kanpur* 85)

PART – V

NUCLEAR PHYSICS

27. Introduction to the Nucleus
28. Nuclear Structure
29. Detectors of Nuclear Radiations
30. Particle Accelerators
31. Radioactivity
32. Radiation Hazards
33. Nuclear Reactions
34. Artificial Transmutation of Elements
35. Nuclear Fission and Fusion
36. Nuclear Reactors
37. Cosmic Rays and the Universe
38. Elementary Particles I
39. Elementary Particles II
40. Interaction of Charged Particles with Matter



CHAPTER 27

INTRODUCTION TO THE NUCLEUS

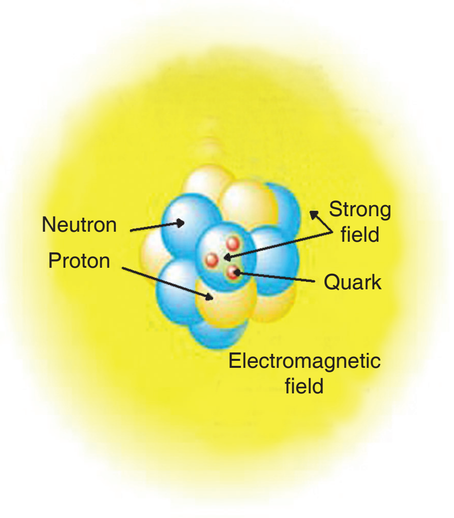
AT A GLANCE

27.1	Introduction	**27.2**	Classification of Nuclei
27.3	General Properties of Nucleus	**27.4**	Binding Energy
27.5	Nuclear Stability	**27.6**	Theories of Nuclear Composition
27.7	Nuclear Forces	**27.8**	Meson Theory of Nuclear Forces
27.9	Models of Nuclear Structure	**27.10**	The Liquid Drop Model
27.11	The Shell Model	**27.12**	The Collective Model
27.13	Determination of Nuclear Radius-Mirror Nuclei Method		
27.14	Nuclear Fission		
27.15	Applications of Semi-Empirical Mass-Formula		

27.1 Introduction

The atomic nucleus was discovered in 1911 by Rutherford. Rutherford's α-particle scattering experiments showed that the atom consists of a very small nucleus (≈ 10–14m in diameter) surrounded by orbiting electrons. It is the purpose of this book to explain all aspects of the nucleus, its structure and its behaviour under various conditions. In this chapter we introduce some of its most basic characteristics, its mass, size, shape, and other externally observable properties. We also consider some deeper questions, such as the force that holds the nucleus together.

Let us begin by reviewing a few fundamental facts that are probably already familiar. All atomic nuclei are made up of elementary particles called *protons* and *neutrons.* A proton has a positive charge of the same



The Nucleus.

magnitude as that of an electron. A neutron is electrically neutral. The proton and the neutron are considered to be two different charge states of the same particle which is called a *nucleon.*

A species of nucleus, known as a *nuclide,* is represented schematically by $_ZX^A$ where Z, the *atomic number* indicates the number of protons, A, the mass number, indicates the total number of protons plus neutrons and X is the chemical symbol of the species.

$$N = \text{Number of neutrons} = A - Z.$$

As an example, the chlorine nucleus $_{17}\text{Cl}^{35}$ has $Z = 17$ protons, $A = 35$ nucleons and $N = 35 - 17 = 18$ neutrons.

27.2 Classification of Nuclei

Atoms of different elements are classified as follows :

(*i*) *Isotopes* are nuclei with the same atomic number Z but different mass numbers A. The nuclei $_{14}\text{Si}^{28}$, $_{14}\text{Si}^{29}$, $_{14}\text{Si}^{30}$, and $_{14}\text{Si}^{32}$ are all isotopes of silicon. The isotopes of an element all contain the same number of protons but have different number of neutrons. Since the nuclear charge is what is ultimately responsible for the characteristic properties of an atom, all the isotopes of an element have identical chemical behaviour and differ physically only in mass.

(*ii*) Those nuclei, with the same mass number A, but different atomic number Z, are called *isobars.* The nuclei $_8O^{16}$ and $_7N^{16}$ are examples of isobars. The isobars are atoms of different elements and have different physical and chemical properties.

(*iii*) Nuclei, with an equal number of neutrons, that is, with the same N, are called *isotones.* Some isotones are $_6C^{14}$, $_7N^{15}$ and $_8O^{16}$ ($N = 8$ in each case).

(*iv*) There are atoms, which have the same Z and same A, but differ from one another in their nuclear energy states and exhibit differences in their internal structure. These nuclei are distinguished by their different life times. Such nuclei are called *isomeric nuclei* or *isomers.*

(*v*) Nuclei, having the same mass number A, but with the proton and neutron number interchanged (that is, the number of protons in one is equal to the number of neutrons in the other) are called *mirror nuclei.*

EXAMPLE. $_4Be^7 (Z = 4 \text{ and } N = 3)$ and $_3Li^7 (Z = 3 \text{ and } N = 4)$.

Note

Nuclei, consisting of an even (odd) number of protons and even (odd) number of neutrons, are said to be *even-even* (odd-odd). Nuclei consisting of an even (odd) number of protons and an odd (even) number of neutrons are said to be *even-odd* (odd-even).

27.3 General Properties of Nucleus

Nuclear size. Rutherford's work on the scattering of α-particles showed that the mean radius of an atomic nucleus is of the order of 10^{-14} to 10^{-15} m while that of the atom is about 10^{-10} m. Thus the nucleus is about 10000 times smaller in radius than the atom.

The empirical formula for the nuclear radius is

$$R = r_o A^{1/3}$$

where A is the mass number and $r_o = 1.3 \times 10^{-15}$ m $= 1.3\,fm$. Nuclei are so small that the fermi (fm) is an appropriate unit of length. $1 fm = 10^{-15}$ m. From this formula we find that the radius of the $_6C^{12}$ nucleus is $R \approx (1.3)\,(12)^{1/3} = 3fm$. Similarly, the radius of the $_{47}Ag^{107}$ nucleus is 6.2 fm and that of the $_{92}U^{238}$ nucleus is 8.1 fm.

The nuclear radius may be estimated from the scattering of neutrons and electrons by the nucleus, or by analysing the effect of the finite size of the nucleus on nuclear and atomic binding energies.

Fast neutrons of about 100 MeV energy, whose wavelength is small compared to the size of the nucleus, are scattered by nuclear targets. The fraction of neutrons scattered at various angles can be used to deduce the nuclear size. The results of these experiments indicate that the radius of a nucleus is given by $R \approx r_0 A^{1/3}$ where $r_0 \approx 1.3 - 1.4\, fm$. The scattering can be done with proton beams as well. In this case, however, the effects due to Coulomb interaction have to be separated out. The observations are in agreement with the equation $R \approx r_0 A^{1/3}$ with $r_0 \approx 1.3 - 1.4\, fm$.

The scattering of fast electrons of energy as high as 10^4 MeV, with a wavelength of about $0.1\, fm$, has the advantage that it can directly measure the charge density inside a nucleus. The results of the experiment are in agreement with the equation $R \approx r_0 A^{1/3}$ but with a somewhat smaller value of $r_0 \approx 1.2\, fm$. The slight difference in the value of r_o may be ascribed to the fact that the electron scattering measures the charge density whereas the neutron and proton scattering experiments measure the region of large nuclear potential, which may be expected to be somewhat larger than the size of the nucleus.

EXAMPLE. *The radius of* Ho^{165} *is* 7.731 *fermi. Deduce the radius of* He^4.

SOL. Let R_1, A_1 and R_2, A_2 be the radius and mass number of Ho^{165} and He^4 respectively. Then $R_1 = r_0 A_1^{1/3}$ and $R_2 = r_0 A_2^{1/3}$.

$$\therefore \qquad \frac{R_1}{R_2} = \frac{A_1^{1/3}}{A_2^{1/3}} \text{ or}$$

$$R_2 = \frac{R_1 A_2^{1/3}}{A_1^{1/3}} = \frac{7.731 \times 4^{1/3}}{(165)^{1/3}} = 2.238\ fm.$$

Nuclear mass. We know that the nucleus consists of protons and neutrons. Then the mass of the nucleus should be

assumed nuclear mass $= Zm_p + Nm_n$

where m_p and m_n are the respective proton and neutron masses and N is the neutron number. Nuclear masses are experimentally measured accurately by mass spectrometers. Measurements by mass spectrometer, however, show that

$$\text{real nuclear mass} < Zm_p + Nm_n.$$

The difference in masses

$$Zm_p + Nm_n - \text{real nuclear mass} = \Delta\, m$$

is called the *mass defect.*

Nuclear density. The nuclear density ρ_N can be calculated from $\rho_N = \dfrac{\text{Nuclear mass}}{\text{Nuclear volume}}$.

Nuclear mass $= Am_N$ where A = mass number and m_N = mass of the nucleon $= 1.67 \times 10^{-27}$ kg.

Nuclear volume $= \dfrac{4}{3}\pi R^3 = \dfrac{4}{3}\pi (r_0 A^{1/3})^3 = \dfrac{4}{3}\pi r_0^3 A$

$$\therefore \qquad \rho_N = \frac{Am_N}{\frac{4}{3}\pi r_0^3 A} = \frac{m_N}{\frac{4}{3}\pi r_0^3} = \frac{(1.67 \times 10^{-27})}{\frac{4}{3}\pi (1.3 \times 10^{-15})^3}$$

$$= 1.816 \times 10^{17} \text{ kg m}^{-3}.$$

Note the high value of the density of the nucleus. This shows that the nuclear matter is in an extremely compressed state. Certain stars (the "*white dwarfs*") are composed of atoms whose electron shells have collapsed owing to enormous pressure, and the densities of such stars approach that of pure nuclear matter.

Nuclear charge. The charge of the neucleus is due to the protons contained in it. Each proton has a positive charge of 1.6×10^{-19} C. The nuclear charge is Ze where Z is the atomic number of the nucleus. The value of Z is known from X-ray scattering experiments, from the nuclear scattering of α-particles, and from the X-ray spectrum.

Spin angular momentum. Both the proton and neutron, like the electron, have an intrinsic spin. *The spin angular momentum* is computed by $L_s = \sqrt{l(l+1)}\ h/2\pi$ where the quantum number l, commonly called the spin, is equal to 1/2. The spin angular momentum, then has a value $L_s = \frac{\sqrt{3}}{2}\frac{h}{2\pi}$.

Resultant angular momentum. In addition to the spin angular momentum, the protons and neutrons in the nucleus have an *orbital angular momentum.* The resultant angular momentum of the nucleus is obtained by adding the spin and orbital angular momenta of all the nucleons within the nucleus. The total angular momentum of a nucleus is given by $L_N = \sqrt{l_N(l_N+1)}\ \frac{h}{2\pi}$. This total angular momentum is called *nuclear spin*.

Nuclear magnetic dipole moments. We know that the spinning electron has an associated magnetic dipole moment of 1 Bohr magneton. *i.e.* $\mu_e = \frac{eh/2\pi}{2m_e}$. Proton has a positive elementary charge and due to its spin, it should have a magnetic dipole moment. According to Dirac's theory, $\mu_N = \frac{eh/2\pi}{2m_p}$ where m_p is the proton mass. μ_N is called a *nuclear magneton* and is the unit of nuclear magnetic moment. μ_N has a value of 5.050×10^{-27} J/T. Since $m_p = 1836\ m_e$, the nuclear magneton is only 1/1836 of a Bohr magneton. For nucleons, however, measurements give $\mu_p = 2.7925\ \mu_N$ and $\mu_n = -1.9128\ \mu_N$. Physicists have found that it is especially hard to understand how the neutron, which is a neutral particle, can have a magnetic moment.

Electric quadrupole moment. In addition to its magnetic moment, a nucleus may have an electric quadrupole moment. An electric dipole moment is zero for atoms and nuclei in stationary states. This is a consequence of the symmetry of nuclei about the centre of mass. However, this symmetry does not need to be spherical; there is nothing precluding the nucleus from assuming the shape of an ellipsoid of rotation, for instance. Indeed most nuclei do assume approximately such a shape, and the deviation from spherical symmetry is expressed by a quantity called the *electric quadrupole moment.* It is defined as

$$Q = \left(\frac{1}{e}\right) \int (3z^2 - r^2)\ \rho\ d\tau$$

where ρ is the charge density in the nucleus. Q is actually a measure of the eccentricity of the ellipsoidal nuclear surface. Evidently $Q = 0$ for a spherically symmetric charge distribution. A charge distribution stretched in the z-direction (prolate) will give a positive quadrupole moment, and an oblate distribution will give a negative quadrupole moment (Fig. 27.1). Since the expression is divided by the electronic charge, the dimension of the quadrupole moment is that of an area. In nuclear physics, area is measured in barns. (1 barn = 10^{-28} m^2).

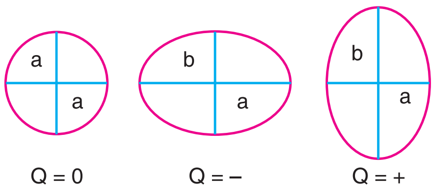
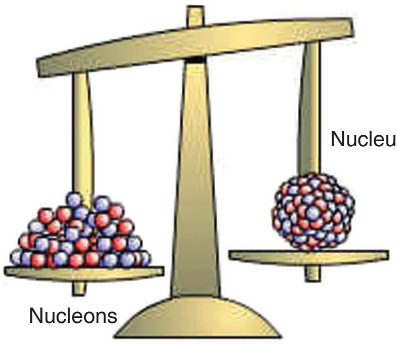


Fig. 27.1

27.4 Binding Energy

The theoretical explanation for the mass defect is based on Einstein's equation $E = mc^2$. When the Z protons and N neutrons combine to make a nucleus, some of the mass ($\Delta\ m$) disappears because it is converted into an amount of energy $\Delta\ E = (\Delta\ m)\ c^2$. This energy is called the *Binding energy* (*B.E*) of the nucleus. To disrupt a stable nucleus into its constituent protons and neutrons, the minimum energy required is the binding energy. The magnitude of the B.E. of a nucleus determines its *stability against disintegration*. If the B.E. is large, the nucleus is stable. A nucleus having the least possible energy, equal to the B.E., is said to be in the *ground state.* If the nucleus has an energy $E > E_{min}$, it is said to be in the *excited state.* The case $E = 0$ corresponds to dissociation of the nucleus into its constituent nucleons.



Binding Energy.

If M is the experimentally determined mass of a nuclide having Z protons and N neutrons,

$$\text{B.E.} = \{(Zm_p + Nm_n) - M\}c^2.$$

If B.E. > 0, the nucleus is stable and energy must be supplied from outside to disrupt it into its constituents. If B.E.< 0, the nucleus is unstabe and it will disintegrate by itself.

Illustration: Let us illustrate the calculation of B.E. by taking the example of the deuteron. A deuteron is formed by a proton and a neutron.

Mass of proton = 1.007276u.

Mass of neutron = 1.008665u.

∴ Mass of proton + neutron in free state = 2.015941u

Mass of deuteron nucleus = 2.013553u

∴ Mass defect = Δm = 0.002388u

B.E. = 0.002388 × 931 = 2.23 MeV (∵ lu = 931 MeV)

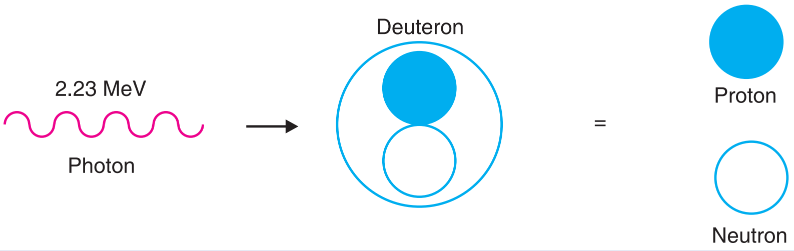


Fig. 27.2

Thus, when a deuteron is formed from a free proton and neutron, 2.23 MeV of energy is liberated. Conversely, 2.23 MeV must be supplied from an external source to break a deuteron up into a proton and a neutron. This is confirmed by experiments that show that a gamma-ray photon with a minimum energy of 2.23 MeV can split a deuteron into a free neutron and a free proton (Fig. 27.2).

EXAMPLE 1. *Calculate the binding energy of an α-particle and express the result both in MeV and joules.*

SOL. Mass of 2 protons + 2 neutrons = (2 × 1.007276 + 2 × 1.008665)u

= 4.031882u.

$$\text{Mass of the } \alpha\text{-particle} = 4.001506u.$$

Mass defect $\Delta m = (4.031882 - 4.001506)\, u = 0.030376\, u$

$\therefore$ $\text{B.E.} = (0.030376 \times 931.3) \text{ MeV} = 28.29 \text{ MeV.}$

$= 45.32 \times 10^{-13}$ J

EXAMPLE 2. *Given the following isotope masses:*

$$_3Li^7 = 7.016004,\ _3Li^6 = 6.015125 \text{ and } _0n^1 = 1.008665\ u$$

Calculate the B.E. of a neutron in the $_3Li^7$ nucleus. Express the result in u, MeV and joules.

SOL. B.E. of a neutron in $u = \text{M}\,(_3Li^6 + {_0n^1}) - \text{M}\,(_3Li^7)$

$= 6.015125 + 1.008665 - 7.016004$

$= 0.007786\, u.$

In MeV this becomes 0.007786 × 931 MeV = 7.35 MeV.

In joules, $7.35 \times 1.6 \times 10^{-13}$ J $= 1.18 \times 10^{-12}$ J.

Stability of Nucleus and Binding Energy

$$\text{B.E per nucleon} = \frac{\text{Total B.E. of a nucleus}}{\text{The number of nucleons it contains}}.$$

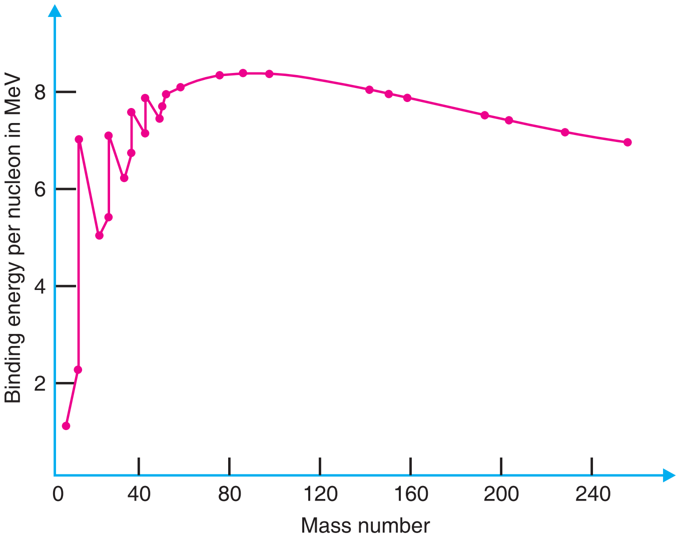


Fig. 27.3

The Binding Energy per nucleon is plotted as a function of mass number A in Fig. 27.3. The curve rises steeply at first and then more gradually until it reaches a maximum of 8.79 *MeV* at $A = 56$, corresponding to the iron nucleus $_{26}Fe^{56}$. The curve then drops slowly to about 7.6 *MeV* at the highest mass numbers. Evidently, nuclei of intermediate mass are the most stable, since the greatest amount of energy must be supplied to liberate each of their nucleons. This fact suggests that a large amount of energy will be liberated if heavier nuclei can somehow be split into lighter ones or if light nuclei can somehow be joined to form heavier ones. The former process is known as *nuclear fission* and the latter as *nuclear fusion.* Both the processes indeed occur under proper circumstances and do evolve energy as predicted.

Packing fraction. The ratio between the mass defect (Δm) and the mass number (A) is called the packing fraction (f);

$$f = \Delta m / A.$$

Packing fraction means the mass defect per nucleon. Since atomic masses are measured relative to C–12, the packing fraction for this isotope is zero. Packing fraction is a measure of the comparative stability of the atom.

Packing fraction is defined as

$$\text{Packing fraction} = \frac{\textit{Isotopic mass} - \textit{Mass number}}{\textit{Mass number}} \times 10^4.$$

Packing fraction may have a *negative* or a *positive* sign. If packing fraction is negative, the isotopic mass is less than the mass number. In such cases, some mass gets transformed into energy in the formation of that nucleus, in accordance with Einsein's equation $E = mc^2$. Such nuclei, therefore, are more stable. A positive packing fraction would imply a tendency towards instability. But this is not quite correct, especially for elements of low atomic masses.

A plot of packing fraction against the corresponding mass numbers of the various elements is shown in Fig. 27.4. It is seen that helium, carbon and oxygen atoms of mass numbers 4, 12 and 16 respectively, do not fall on this curve. Their paking fractions have small values. These elements are, therefore, stable. The transition elements, with mass numbers in the neighbourhood of 45, have lowest packing fractions with a negative sign, which indicates their high stability. The packing fraction beyond mass number 200 becomes positive and increases with increase in mass number. This indicates increasing instability of these elements. Elements with mass numbers beyond 230 are radioactive and undergo disintegration spontaneously.

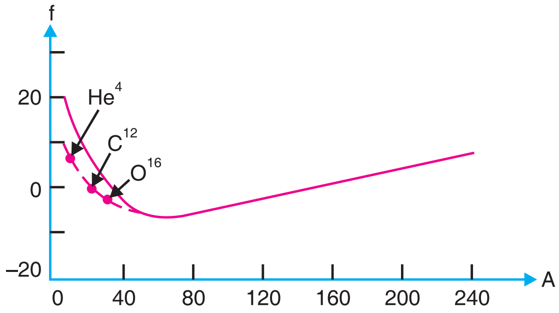


Fig. 27.4

27.5 Nuclear Stability

Table 27.1 shows how the 272 stable nuclei found in nature are classified according to even and odd numbers of protons and neutrons.

TABLE 27.1

Protons	*Neutrons*	*Stable Nuclides*
even	even	160
even	odd	56
odd	even	52
odd	odd	4
		272

The combination of an even number of protons and an even number of neutrons, composing the nucleus, is evidently preferred by nature for stable nuclides. The odd-odd combination of stable nuclides is found only in the light elements. The number of even-odd combinations is about the same. A plot of the number of neutrons versus the number of protons for the stable nuclides is shown in Fig. 27.5. Notice that for $Z < 20$, the stability line is a straight line with $Z = N$. For the heavier nuclides $Z > 20$, $N > 20$, the stability curve bends in the direction of $N > Z$. For example $_{20}Ca^{48}$ has $N = 28$, $Z = 20$; for larger values of Z, the tendency is more pronounced, as in the case of $_{91}Pa^{232}$ which has $N = 141$, $Z = 91$.

Evidently, for large values of Z, the coulomb electrostatic repulsion becomes important, and the number of neutrons must be greater to compensate this repulsive effect.

Thus the curve of Fig. 27.5 departs more and more from the $N = Z$ line as Z increases. For maximum stability, there is an optimum value of neutron/proton ratio. The number of neutrons N $(= A - Z)$ required for maximum stability is plotted as a function of proton number Z in Fig. 27.6. All the stable nuclei fall within the shaded region. Nuclei above and below the shaded region are unstable. Artificial radioactive nuclei lie at the fringe of the region of stability. All nuclei with

Z > 83, and A > 209 spontaneously transform themselves into lighter ones through the emission of α and β particles. α and β decays enable an unstable nucleus to reach a stable configuration.

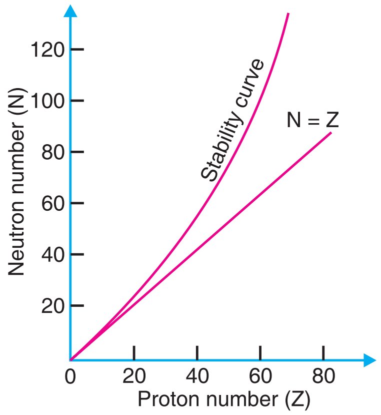


Fig. 27.5

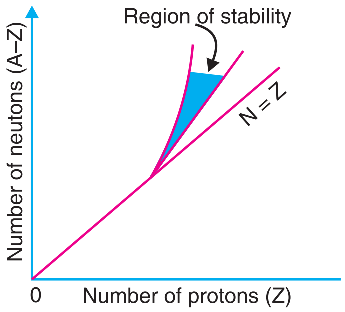


Fig. 27.6

27.6 Theories of Nuclear Composition

Several theories were proposed about the nuclear constitution. Prominent among these are proton–electron theory and proton–neutron theory.

Proton-electron hypothesis. Before the discovery of neutron by Chadwick (1932), it was assumed that the nucleus consists of protons and electrons. This concept arose because electrons were found to be emitted by radioactive nuclei in β-decay process. According to this hypothesis, a nucleus of mass number *A* and atomic number *Z* contains *A* protons and *A–Z* electrons. The nucleus is surrounded by *Z* electrons, so that the atom is electrically neutral.

Example. *The nucleus of* ${}_{11}Na^{23}$ *should consist of 23 protons and 12 electrons. The total nuclear charge is 11 which is balanced by —11 units of electron charge in the extranuclear space.*

There was very strong theoretical evidence for the fact that electrons cannot exist in the nucleus. We discuss below some very strong arguments which forbid the existence of electrons inside the nucleus.

Why electrons cannot be present inside the nucleus?

***(i)* Nuclear size.** Typical nuclei are less than 10^{-14} m in radius. If an electron exists inside a nucleus, the uncertainty in its position (Δx) may not exceed 10^{-14} m. According to Heisenberg's uncertainty principle, uncertainty in the electron's momentum is

$$\Delta p \ \geq \ \frac{h/2\pi}{\Delta x} \geq \frac{1.054 \times 10^{-34}}{10^{-14}} \geq 1.1 \times 10^{-20} \text{ kg ms}^{-1}.$$

If this is the uncertainty in the electron's momentum, the momentum itself must be at least comparable in magnitude.

∴ Approximate momentum of the electron = p = 1.1×10^{-20} kg ms^{-1}. An electron whose momentum is 1.1×10^{-20} kg ms^{-1} has a K.E. (T) many times greater than its rest energy $m_0 c^2$ *i.e.*, $T >> m_0 c^2$. Hence we can use the extreme relativistic formula $T = pc$ to find T.

$$\therefore \qquad T = (1.1 \times 10^{-20})\,(3 \times 10^{8}) = 3.3 \times 10^{-12}\,J$$

$$= \frac{3.3 \times 10^{-12}}{1.6 \times 10^{-13}}\,MeV = 20.63\,MeV\cdot$$

This shows that if an electron exists in the nucleus, the K.E. of the electron must be more than 20 *MeV*. Electrons of such large energy are never found to be emitted during β-decay. The maximum energy of a β-particle emitted is only 2 to 3 *MeV*. Hence we conclude that electrons cannot be present within nuclei.

(*ii*) Nuclear spin. Electrons and protons have a spin of 1/2. Thus nuclei with an even number of protons and electrons should have integral spins, while those with an odd number of protons and electrons should have half-integral spins. Let us consider deuteron as an example. Deuteron nucleus has 3 particles (two protons and one electron). Hence the nuclear spin of deuteron should be 1/2 or 3/2. But experiment shows that the spin of the deuteron is 1. Thus the experimental result is in contradiction to the hypothesis.

(*iii*) Magnetic moment. Protons and electrons are endowed with magnetic properties. The magnetic moment of an electron is about one thousand times that of a proton. If electrons exist inside the nucleus, the magnetic moment of electrons will have a dominating influence and so nuclear magnetic moments ought to be of the same order of magnitude as that of the electron. However, the observed magnetic moments of nuclei are comparable with that of the proton. This experimental fact goes against the electrons existing inside the nucleus.

Due to these reasons, it is concluded that electrons cannot exist in the nucleus. Hence the proton-electron hypothesis regarding the constitution of the nucleus has been given up.

Proton-Neutron hypothesis. After the discovery of neutron by Chadwick in 1932, proton-neutron theory has been gaining support. According to this, a nucleus of atomic number Z and mass number A consists of Z protons and $A–Z$ neutrons. The number of electrons in the extra nuclear space is also Z, so that the atom as a whole is neutral.

Merits. (*i*) This assumption is able to explain the observed value of nuclear spin and nuclear magnetic moment.

(*ii*) There is no difficulty in explaining the existence of isotopes on this hypothesis. Isotopes of a given element differ only in the number of neutrons they contain.

(*iii*) The β-*ray emission of radioactive elements is explained as follows*: *The electron does not pre-exist in the nucleus. The electron is formed just at the instant of emission*, *caused by the transformation of a neutron into a proton. i.e.*, $n \to p + e^-$. *The positron emission is due to the converse process*, *viz*., *when a proton transforms itself into a neutron. i.e*., $p \to n + e^+$.

(*iv*) The emission of α-particles from the nuclei of radioactive elements is due to the combination of 2 protons and 2 neutrons at the instant of emission. Thus, this hypothesis explains β-decay and α-decay.

27.7 Nuclear Forces

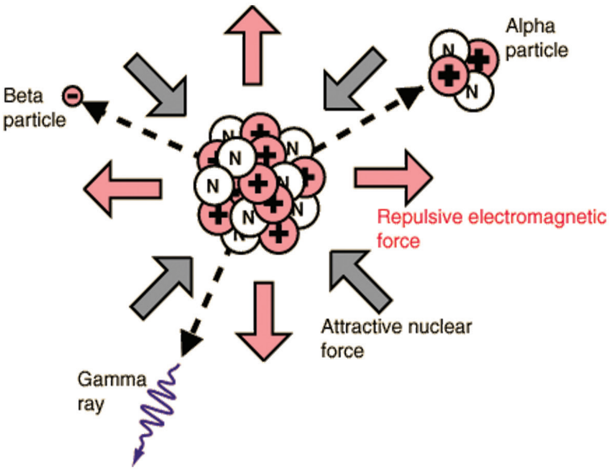
Since stable nuclei exist, it follows that there must be certain forces acting between their nucleons that bind them into the nucleus. These are called *nuclear forces. The nuclear force must be strongly attractive*, in order to overcome the electrostatic repulsion between protons. Only three kinds of attractive forces can be conceived in the nucleus, viz., neutron-neutron (*n-n*), neutron-proton (*n-p*) and proton-proton (*p-p*) interactions.

Yukawa attributed the following characteristics to the nuclear forces.

(1) Nuclear forces are effective only at short ranges. Nuclear forces are appreciable only when the distance between nucleons is of the order of 10^{-15} m or less. The force vanishes for all practical purposes at distances greater than a few times 10^{-15} m. These distances are called the *action radii* or *range* of the nuclear forces. In the up-to-date version of the exchange theory of nuclear forces, it is supposed that interaction between nucleons is accomplished by the exchange of π-mesons. The exchange version of nuclear forces explains their short range action. Let m be the rest mass of the π-meson. The rest energy of the π-meson = $\Delta E = mc^2$. According to Heisenberg's uncertainty principle, the time required for nucleons to exchange π-mesons cannot exceed $\Delta\, t$, for which $\Delta E \; \Delta t \geq {}^{h}\!/_{2}\, \pi$. The distance that a π-meson can travel away from a nucleon in the nucleus during the time $\Delta\, t$, even at a velocity $\approx c$, is

$$R_0 \approx \frac{h/2\pi}{mc} \approx 1.2 \times 10^{-15}\ m.$$

This approximately coincides with the value of the nuclear radius and is of the order of magnitude of the nuclear force range.



Nuclear Forces.

(2) Nuclear forces are charge independent. The nuclear forces acting between two protons, or between two neutrons, or between a proton and a neutron, are the same. It follows that nuclear forces are of a non-electric nature.

(3) Nuclear forces are the strongest known forces in nature.

(4) Nuclear forces have saturation property. Nuclear forces are limited in range. As a result, each nucleon interacts with only a limited number of nucleons nearest to it. This effect is referred to as the *saturation* of nuclear forces.

27.8 Meson Theory of Nuclear Forces

According to the meson theory of nuclear forces, all nucleons consist of identical cores surrounded by a "*cloud*" of one or more mesons. Mesons may be neutral or may have a positive or negative charge. The sole difference between neutrons and protons is supposed to lie in the composition of their respective meson clouds. Yukawa assumed that π meson is exchanged between the nucleons and that this exchange is responsible for the nuclear binding forces. The forces that act, between one neutron and another, and between one proton and another, are the result of the exchange of neutral mesons (π°) between them. The force between a neutron and a proton is the result of the exchange of charged mesons (π^+ and π^-) between them.

Thus a neutron emits a π^- meson and is converted into a proton:

$$n \rightarrow p + \pi^-$$

The absorption of the π^- meson by the proton (with which the neutron was interacting) converts it into a neutron:

$$p + \pi^- \rightarrow n.$$

In the reverse process, a proton emits a π^+ meson, becoming a neutron and the neutron, on receiving this π^+ meson, becomes a proton :

$$p \rightarrow \text{n} + \pi+$$

$$n + \pi^+ \rightarrow p.$$

Thus in the nucleus of an atom, attractive forces exist between (1) proton and proton (2) proton and neutron and (3) neutron and neutron. These forces of attraction are much larger than the electrostatic force of repulsion between the protons, thus giving a stability to the nucleus.

Just as a photon is a quantum of electromagnetic field, a meson is a quantum of nuclear field. Yukawa considered the equation for particle of mass m as,

$$\left(\nabla^2 - \frac{m^2c^2}{(h/2\pi)^2} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\phi = 0 \qquad \text{...(1)}$$

This is a relativistic equation valid only for spinless particles.

Separating the time dependent part, the equation for the radial part is

$$(\nabla^2 - \mu^2)\,\phi\,(r) = 0 \qquad \text{...(2)}$$

where $\mu = mc/\,(h/2\pi)$.

The solution of Eq. (2) is $\phi\,(r) = -g\,\dfrac{e^{-\mu r}}{r}$ ...(3)

Here g is a constant, which plays the same role as the charge q in electromagnetic theory. In analogy with electromagnetism, the potential between two nucleons is then given by

$$V(r) = -g^2\,\frac{e^{-\mu r}}{r} \qquad \text{...(4)}$$

Here g^2 is called the '*Coupling constant*'. This argument made Yukawa predict the existence of *pion* as a quantum of nuclear force field.

The range of the pion field is $\dfrac{h/2\pi}{m_\pi\,c} \approx 1.4\ fm.$

The form of $V\,(r)$ given by Eq. (4) is known as the *one-pion-exchange potential* (*OPEP*).

On the basis of the range of nuclear force and the uncertainty principle, it is possible to estimate the mass of the meson. According to uncertainty principle $\Delta E \times \Delta t = h/2\pi$ where ΔE and Δt are the uncertainties in energy and time. The range of nuclear force is $R \approx 1.4 \times 10^{-15}\ m$. Let us assume that the meson travels between nuclei at approximately the speed of light c. Let Δt be the time interval between the emission of meson from one nucleon and the absorption by the other nucleon.

Then $\Delta t = R/c.$ $\quad\therefore\quad \Delta E = \dfrac{(h/2\pi)}{\Delta t}$

$\therefore$ The minimum meson mass is specified by $m \geq \dfrac{(h/2\pi)}{Rc}$.

In terms of the electronic mass m_e, the mass of the meson is

$$\frac{m}{m_e} = \frac{h/2\pi}{m_e\ Rc} = \frac{1.054 \times 10^{-34}}{(9.108 \times 10^{-31})\,(1.4 \times 10^{-15})\,(3 \times 10^8)} = 275$$

i.e., mass of the meson ≈ 275 × mass of electron.

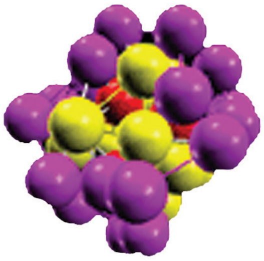
In 1947, Powell discovered π meson of mass about 273 m_e. This particle showed strong interaction with nucleons and was recognized as the Yukawa particle.

The discovery of the meson of mass of about 273 electron mass and the existence of positive, negative and neutral mesons, lends some support to this theory. The experimental values of the magnetic moments of a free proton and of a free neutron also lend some support to the "*Yukawa's meson field theory*" of nuclear forces. A free proton is, for a part of its life-time, a neutron with a closely bound meson. Hence the magnetic moment of a free proton can be the resultant of the true magnetic moment of the proton and the magnetic moment of the meson. Thus the net magnetic moment of a free proton will exceed that given by the simple theory. Similarly, a neutron is for a fraction of its life-time dissociated into a proton and a negative meson. This combination will have a negative magnetic moment. It follows that, though uncharged, a neutron will have a negative magnetic moment.

27.9 Models of Nuclear Structure

The precise nature of the forces acting in the nucleus is unknown. Hence, nuclear models are resorted to for investigation and theoretical prediction of its properties. Such models may be based on (*i*) the extrinsic analogy between the properties of atomic nuclei and those of a liquid drop (*ii*) the electron shell of an atom etc. The corresponding models are called the liquid-drop model, shell model, etc.

27.10 The Liquid Drop Model



Liquid Drop Model.

In the liquid-drop model, the forces acting in the nucleus are assumed to be analogical to the molecular forces in a droplet of some liquid. This model was proposed by Neils Bohr who observed that there are certain marked similarities between an atomic nucleus and a liquid drop. The similarities between the nucleus and a liquid drop are the following:

(*i*) The nucleus is supposed to be spherical in shape in the stable state, just as a liquid drop is spherical due to the symmetrical surface tension forces.

(*ii*) The force of surface tension acts on the surface of the liquid-drop. Similarly, there is a potential barrier at the surface of the nucleus.

(*iii*) The density of a liquid-drop is independent of its volume. Similarly, the density of the nucleus is independent of its volume.

(*iv*) The intermolecular forces in a liquid are short range forces. The molecules in a liquid drop interact only with their immediate neighbours. Similarly, the nuclear forces are short range forces. Nucleons in the nucleus also interact only with their immediate neighbours. This leads to the saturation in the nuclear forces and a constant binding energy per nucleon.

(*v*) The molecules evaporate from a liquid drop on raising the temperature of the liquid due to their increased energy of thermal agitation. Similarly, when energy is given to a nucleus by bombarding it with nuclear projectiles, a compound nucleus is formed which emits nuclear radiations almost immediately.

(*vi*) When a small drop of liquid is allowed to oscillate, it breaks up into two smaller drops of equal size. The process of nuclear fission is similar and the nucleus breaks up into two smaller nuclei.

Semi-empirical mass formula. The liquid-drop model can be used to obtain an expression for the binding energy of the nucleus. Weizacker proposed the semi-empirical nuclear binding energy formula for a nucleus of mass number A, containing Z protons and N neutrons. It is written as

$$\text{B.E.} = aA - bA^{2/3} - \frac{cZ(Z-1)}{A^{1/3}} - \frac{d(N-Z)^2}{A} \pm \frac{\delta}{A^{3/4}}$$

where a, b, c, d and δ are constants.

Explanation of the terms. (1) The first term is called the *volume energy* of a nucleus ($E_v = aA$). The larger the total number of nucleons A, the more difficult it will be to remove the individual protons and neutrons from the nucleus. The B. E. is directly proportional to the total number of nucleons A.

(2) The nucleons, at the surface of the nucleus, are not completely surrounded by other nucleons. Hence energy of the nucleon on the surface is less than that in the interior. The number of surface nucleons depends upon the surface area of the nucleus. A nucleus of radius R has an area of $4\pi R^2 = 4\pi r_o^2 A^{2/3}$. Hence the *surface effect* reduces the B.E. by $E_s = bA^{2/3}$. The negative energy E_s

is called the *surface energy* of a nucleus. It is most significant for the lighter nuclei, since a greater fraction of their nucleons are on the surface.

(3) The *electrostatic repulsion* between each pair of protons in a nucleus also contributes towards decreasing its B.E. The Coulomb energy E_c of a nucleus is the work that must be done to bring together Z protons from infinity into a volume equal to that of the nucleus. Hence $E_c \propto Z(Z-1)/2$ (the number of proton pairs in a nucleus containing Z protons) and E_c is inversely proportional to the nuclear radius $R = r_o A^{1/3}$. E_c is negative because it arises from a force that opposes nuclear stability.

(4) The fourth term $E_a = \dfrac{d\,(N-Z)^2}{A}$ originates from the *lack of symmetry* between the number of protons (Z) and the number of neutrons (N) in the nucleus. The maximum stability of a nucleus occurs when $N = Z$. Any departure from this introduces an asymmetry N–Z, which results in a decrease in stability. The decrease in the B.E. arising from this is called the asymmetric energy (E_a). This is also negative.

(5) The final correction term δ allows for the fact that *even-even* nuclei are more stable than *odd-odd* nuclei. δ is positive for even-even nuclei, is negative for odd-odd nuclei and δ = 0 for an odd A.

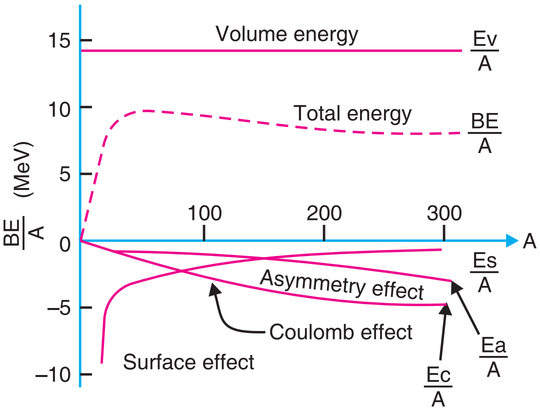


Fig. 27.7

The best values of the constants, expressed in MeV, are $a = 15.760$; $b = 17.810$, $c = 0.711$, $d = 23.702$, $\delta = 34$.

The contributions of the various effects in Weizacker's empirical formula are represented schematically in the graph of Fig. 27.7.

Merits. (1) The liquid drop model accounts for many of the salient features of nuclear matter, such as the observed binding energies of nuclei and their stability against α and β disintegration as well as nuclear fission.

(2) The calculation of atomic masses and binding energies can be done with good accuracy with the liquid drop-model.

However, this model fails to explain other properties, in particular the magic numbers. It fails to explain the measured spins and magnetic moments of nuclei.

EXAMPLE. *Calculate the atomic number of the most stable nucleus for a given mass number A.*

SOL. The most stable nucleus with a given mass number A is that which has the maximum value of the B.E. Thus we have to compute $\dfrac{\partial\,(B.E.)}{\partial\, Z}$ with A constant, and equate it to zero. In the formula for B.E., we can write $Z(Z-1) = Z^2$ and $N - Z = A - 2Z$. Then

$$B.E. = aA - bA^{2/3} - \frac{cZ^2}{A^{1/3}} - \frac{d\,(A-2Z)^2}{A} \pm \frac{\delta}{A^{3/4}}$$

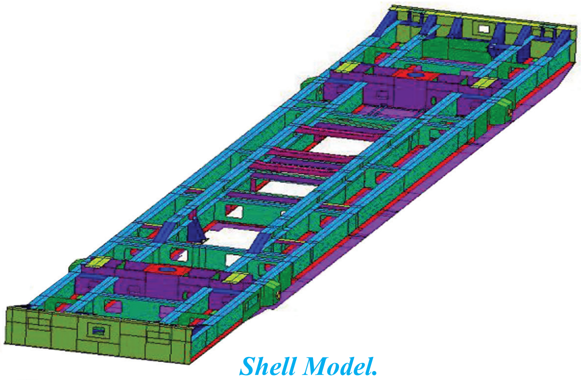
$$\frac{\delta E_b}{\delta Z} = -\,cZA^{-1/3} + 4d\,(A-2Z)\,A^{-1} = 0$$

or, introducing the numerical values of c and d, we have

$$Z = \frac{A}{2 + 0.0157\ A^{2/3}}.$$

For light nuclei, having small A, we can neglect the second term in the denominator. $\therefore\ Z \approx A/2$. This result is confirmed *experimentally*.

27.11 The Shell Model



Shell Model.

The shell model of the nucleus assumes that the energy structure (energy levels of the nucleons) of the nucleus is similar to that of an electron shell in an atom. According to this model, the protons and neutrons are grouped in shells in the nucleus, similar to extra-nuclear electrons in various shells outside the nucleus. The shells are regarded as "filled" when they contain a specific number of protons or neutrons or both. The number of nucleons in each shell is limited by the *Pauli exclusion principle*. The shell model is sometimes referred to as the *independent particle model* because it assumes that each nucleon moves independently of all the other nucleons and is acted on by an average nuclear field produced by the action of all the other nucleons.

Evidence for shell model. It is known that a nucleus is stable if it has a certain definite number of either protons or neutrons. These numbers are known as *magic numbers*. The magic numbers are 2, 8, 20, 50, 82 and 126. Thus nuclei containing 2, 8, 20, 50, 82 and 126 nucleons of the same kind form some sort of closed nuclear shell structures. The main points in favour of this inference are:

(*i*) The inert gases with closed electron shells exhibit a high degree of chemical stability. Similarly, nuclides whose nuclei contain a magic number of nucleons of the same kind exhibit more than average stability.

(*ii*) Isotopes of elements having an isotopic abundance greater than 60% belong to the magic number category.

(*iii*) Tin ($_{50}$Sn) has ten stable isotopes, while calcium ($_{20}$Ca^{40}) has six stable isotopes. So elements with Z = 50, 20 are more than usually stable.

(*iv*) The three main radioactive series (viz., the uranium series, actinium series and thorium series) decay to $_{82}Pb^{208}$ with $Z = 82$ and $N = 126$. Thus lead $_{82}Pb^{208}$ is the most stable isotope. This again shows that the numbers 82 and 126 indicate stability.

(*v*) It has been found that nuclei having a number of neutrons equal to the magic number, cannot capture a neutron because the shells are closed and they cannot contain an extra neutron.

(*vi*) It is found that some isotopes are spontaneous neutron emitters when excited above the nucleon binding energy by a preceding β-decay. These are $_8O^{17}$, $_{36}Kr^{87}$ and $_{54}Xe^{137}$ for which $N = 9$, 51 and 83 which can be written as 8 + 1, 50 +1, and 82 + 1. If we interpret this loosely bound neutron, as a valency neutron, the neutron numbers 8, 50, 82 represent greater stability than other neutron numbers.

It is apparant from the above conclusions that nuclear behaviour is often determined by the excess or deficiency of nucleons with respect to closed shells of nucleons corresponding to the magic numbers. It was, therefore, suggested that nucleons revolve inside the nucleus just as electrons revolve outside in specific permitted orbits. The protons and neutrons move in two separate systems of orbits round the centre of mass of all the nucleons. The extra-nuclear electrons revolve in the Coulomb field of a relatively distant heavy nucleus. But the nucleons move in orbits around a common centre of gravity of all the constituents of the nucleus. Each nucleon shell has a specific maximum capacity. When the shells are filled to capacity, they give rise to particular numbers (the magic numbers) characteristic of unusual stability.

The shell model is able to account for several nuclear phenomena in addition to magic numbers.

(*i*) It is observed that even-even nuclei are, in general, more stable than odd-odd nuclei. This is obvious from the shell model. According to Pauli's principle, a single energy sublevel can have a maximum of two nucleons (one with spin up and other with spin down). Therefore, in an even-even nucleus only *completed* sublevels are present which means greater stability. On the other hand, an odd-odd nucleus contains incomplete sublevels for both kinds of nucleon which means lesser stability.

(*ii*) The shell model is able to predict the total angular momenta of nuclei. In even-even nuclei, all the protons and neutrons should pair off so as to cancel out one another's spin and orbital angular momenta. Thus even-even nuclei ought to have zero nuclear angular momenta, as observed. In even-odd and odd-even nuclei, the half-integral spin of the single "extra" nucleon should be combined with the integral angular momentum of the rest of nucleus for a half-integral total angular momentum. Odd-odd nuclei each have an extra neutron and an extra proton whose half-integral spins should yield integral total angular momenta. Both these predictions are experimentally confirmed.

27.12 The Collective Model

The collective model was proposed by A. Bohr, B. R. Mottleson and James Rainwater. The model combines the best features of the liquid drop model and the shell model. In this model it is assumed that the particles within the nucleus exert a centrifugal pressure on the surface of the nucleus. That results in the permanent deformation to non-spherical shape. As a result, the surface may undergo periodic oscillations.

The particles within the nucleus move in a non-spherical potential. Thus the nuclear distortion reacts on the particles and modifies the independent particle aspect. Thus the nucleus is considered as a shell structure capable of performing oscillations in shape and size. Thus the model can easily describe the drop like properties such as nuclear fission and at the same time it can retain the shell model characteristics. It is capable of explaining not only the large electric quadrupole moments but it can also predict the fine structure of nuclear level spectrum due to the energies associated with the vibrational and rotational motion of the core.

The total energy is expressed as

$$W = E_{rot} + E_{vib} + E_n$$

where E_{rot} is the energy due to rotational motion of the core, E_{vib} is the energy due to the vibrational coordinates and E_n is the energy due to the nucleonic coordinates. So the wave function is the product of three wave functions each containing the respective coordinates.

According to the collective model, all even-even nuclei such as ${}_2\text{He}^4$, ${}_8\text{O}^{16}$, ${}_{20}\text{Ca}^{40}$, ${}_{82}\text{Pb}^{208}$ will have spherical shapes and zero electric quadrupole moment, while the even-odd, odd-even or odd-odd nuclei will have non-spherical shapes and finite electric quardrupole moment. The nuclear energy levels predicted by the model agree closely with the ones given by γ–ray spectra of the nuclei.

27.13 Determination of Nuclear Radius–Mirror Nuclei Method

Two nuclei having same number of nucleons (protons + neutrons) but the number of protons in one of them being equal to the number of neutrons in the other are called the mirror-nuclei (Fig. 27.8.)

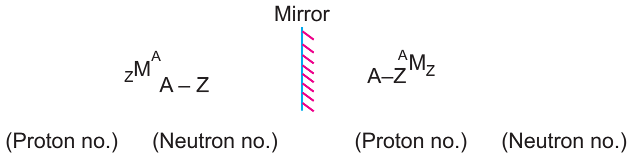


Fig. 27.8

EXAMPLE. *$_6C^{11}$ and $_5B^{11}$ are mirror nuclei. $_6C^{11}$ contains 6 protons and 5 neutrons while $_5B^{11}$ contains 5 protons and 6 neutrons.*

SOL. The other examples of mirror nuclei are: $({}_1^4H, {}_2^3He)$; $({}_3^7Li, {}_4^7Be)$; $({}_6^{13}C, {}_7^{13}N)$; $({}_{19}^{39}K, {}_{20}^{39}Ca)$.

Bethe suggested that the difference of Coulomb energies between two neighbouring mirror nuclei can be used to calculate nuclear radii.

Consider a spherical nucleus of radius R and atomic number Z (Fig. 27.9). Each proton carries a charge $+e$. Thus the total charge is $+Ze$.

$$\therefore \text{ Charge density,} \qquad \rho = \frac{\text{Charge on nucleus}}{\text{Volume of nucleus}} = \frac{Ze}{\frac{4}{3}\pi R^3} \qquad \text{...(1)}$$

The electrostatic energy of nucleus is the work done against electrostatic forces in assembling such a sphere.

Suppose at any instant we have a spherical core of radius x.

$$\text{Charge on this spherical core} = q_1 = \frac{4}{3}\pi x^3 \rho$$

Now we bring a thin spherical shell of radius x and thickness dx from infinity in the process of assembling the charge.

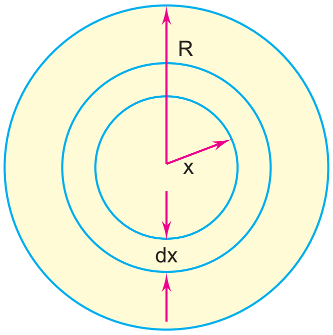


Fig. 27.9

The charge on this shell is

$$q_2 = \text{volume of the shell} \times \text{charge} \times \text{density} = (4\pi x^2\, dx)\, \rho.$$

The electrostatic potential energy of this spherical core and thin spherical shell

$$dE_c = \frac{1}{4\pi\epsilon_0} \frac{q_1 q_2}{x}$$

$$= \frac{1}{4\pi\epsilon_0} \frac{\left(\frac{4}{3}\pi x^3 \rho\right)\left(4\pi x^2\, dx\, \rho\right)}{x} = \frac{1}{4\pi\epsilon_0}\left(\frac{16}{3}\pi^2\rho^2 x^4 dx\right)$$

Total electrostatic potential energy of nucleus is obtained by integrating this expression between limits 0 and R.

$$E_c = \int_0^R \frac{1}{4\pi\epsilon_0}\left(\frac{16}{3}\pi^2\rho^2 x^4 dx\right) = \frac{1}{4\pi\epsilon_0}\frac{16}{3}\pi^2\rho^2\left[\frac{R^5}{5}\right]$$

Substituting value of ρ from Eq. (1), we get

$$E_c = \frac{1}{4\pi\in_0}\frac{16}{3}\pi^2 \cdot \left(\frac{Ze}{\frac{4}{3}\pi R^3}\right)^2 \cdot \frac{R^5}{5} = \frac{1}{4\pi\in_0}\cdot\frac{3}{5}\cdot\frac{Z^2e^2}{R} \qquad ...(2)$$

The difference in Coulomb energies between two neighbouring nuclei of charges (Ze) and $(Z+1)e$ having the same radius R is

$$\Delta E_c = \frac{1}{4\pi\in_0}\frac{3}{5}\frac{(Z+1)^2e^2}{R} - \frac{1}{4\pi\in_0}\frac{3}{5}\frac{Z^2e^2}{R} = \frac{1}{4\pi\in_0}\frac{3}{5}\frac{e^2}{R}\left[(Z+1)^2 - Z^2\right]$$

$$\Delta E_c = \frac{1}{4\pi\in_0}\frac{3}{5}\frac{e^2}{R}(2Z+1) \qquad ...(3)$$

In the mirror nuclei $(2Z+1) = A$, where A is mass number. Eq. (3) becomes

$$\Delta E_c = \frac{1}{4\pi\in_0}\frac{3}{5}\frac{e^2 A}{R} \qquad ...(4)$$

Taking experimental observations of mirror nuclei (${}_{14}Si^{29}$, ${}_{15}P^{29}$), the difference in Coulomb-energies comes out to be 4.96 MeV. Substituting this value, $R = 4.96$ fermi.

From the relation $R = R_0 A^{1/3}$, we obtain the radial parameter, $R_0 \approx 1.5$ fermi.

27.14 Nuclear Fission

EXAMPLE 1. *Obtain an expression for energy released in symmetric fission.*

(*Mumbai University, October 2010*)

SOL. In a *'symmetric fission'*, a nucleus (Z, A) breaks into two equal halves ($Z/2$, $A/2$). Semi-empirical mass formula gives a satisfactory account of fission energetics.

The energy released in a symmetric fission is:

$$Q = M(Z, A) - 2M(Z/2, A/2) \qquad ...(1)$$

From the semi-empirical mass formula, binding energy for a nucleus (Z, A) is:

$$B(Z, A) = a_v\, A - a_s\, A^{2/3} - a_c \frac{Z(Z-1)}{A^{1/3}} - a_a \frac{(A-2Z)^2}{A} - a_p\, A^{-3/4} \qquad ...(2)$$

Binding energy for two fission fragments ($Z/2$, $A/2$) is:

$$2B(Z/2, A/2) = a_v\, A - 2a_s\left(\frac{A}{2}\right)^{2/3} - 2a_c \frac{\frac{Z}{2}\left(\frac{Z}{2}-1\right)}{\left(\frac{A}{2}\right)^{1/3}} - 2a_a \frac{(A/2-Z)^2}{A/2} - 2a_p\left(\frac{A}{2}\right)^{-3/4} \qquad ...(3)$$

$$Q = a_s\, A^{2/3}\left(1 - 2^{1/3}\right) + a_c \frac{Z(Z-1)}{A^{1/3}}\left(1 - \frac{1}{2^{2/3}}\right) + \text{pairing energy difference.}$$

The difference in pairing energy is small and thus can be neglected as an approximation.

$$\therefore \qquad Q = A^{2/3}\left(-0.260 a_s + 0.370 \frac{Z(Z-1)}{A} a_c\right) \qquad ...(4)$$

Using the values of $a_s = 13$ MeV and $a_c = 0.60$ MeV, we get

$$Q \sim A^{2/3}\left(-3.38\text{ MeV}+0.22\text{ MeV}\,\frac{Z(Z-1)}{A}\right) \qquad ...(5)$$

From this we see that Q becomes > 0 for

$$\frac{Z^2}{A} \geq \frac{3.38}{0.22} \sim 15 \qquad ...(6)$$

Thus for all nuclei having, $Z > 35$, $A > 80$, fission is possible and would release energy.

For the symmetric fission of ${}^{238}_{92}U$,

$$Q = (238)^{2/3}\left(-3.38+0.22\,\frac{92\,(92-1)}{238}\right)\text{MeV} \sim 170\text{ MeV} \qquad ...(7)$$

EXAMPLE 2. *Explain mass yield in asymmetrical fission and discuss mass distribution of fission fragments for different fission chains of U^{235}.* (*Mumbai, 2010*)

SOL. Consider the following fission reaction:

$$\underset{\text{Compound nucleus}}{{}^{235}_{92}\text{U} + {}^{1}_{0}n \rightarrow \left[{}^{236}_{92}\text{U}\right]^*} \rightarrow {}^{143}_{56}\text{Ba} + {}^{90}_{36}\text{Kr} + 3\left({}^{1}_{0}n\right) + Q \qquad ...(1)$$

The stable isotope of Barium with maximum A is ^{138}Ba. The stable isotope of Krypton with maximum A is ^{86}Kr. To get rid of excess neutrons, these fission fragments must exhibit much beta activity. Experimentally this is confirmed by the following beta decay observations:

$$\begin{tikzcd}[column sep=small]
{}^{143}_{86}\text{Ba} \arrow[r] & {}^{143}_{57}\text{La} + \beta^- + \bar{\nu} \arrow[dr] & & & \\
& & {}^{143}_{58}\text{Ce} + \beta^- + \bar{\nu} \arrow[dr] & & \\
& & & {}^{143}_{59}\text{Pr} + \beta^- + \bar{\nu} \arrow[dr] & \\
& & & & {}^{143}_{60}\text{Nd} + \beta^- + \bar{\nu} \\
& & & & \text{(Stable)}
\end{tikzcd} \qquad ...(2)$$

$$\begin{tikzcd}[column sep=small]
{}^{90}_{36}\text{Kr} \arrow[r] & {}^{90}_{37}\text{Rb} + \beta^- + \bar{\nu} \arrow[dr] & & & \\
& & {}^{90}_{38}\text{Sr} + \beta^- + \bar{\nu} \arrow[dr] & & \\
& & & {}^{90}_{39}\text{Y} + \beta^- + \bar{\nu} \arrow[dr] & \\
& & & & {}^{90}_{40}\text{Zr} + \beta^- + \bar{\nu} \\
& & & & \text{(Stable)}
\end{tikzcd} \qquad ...(3)$$

Eq. (1) is only one of the many ways in which ${}^{235}_{92}U$ -nucleus undergoes fission.

There are more than 30 modes of fission, in each of which a different pair of nuclei (fission-fragments) is formed. Eqs. (2) and (3) are examples of a *fission chain.* There are about 60 such chains known experimentally.

Mass Distribution of Fission Fragments for Different Fission Chains of ${}^{235}_{92}$U

The fission yield $Y(A)$ is defined as the ratio of number of nuclei of mass number A formed in fission to the total number of fissions, *i.e.,*

$$Y(A) = \frac{\text{Number of nuclei of mass number } A \text{ formed in fission } (N_A)}{\text{Total number of fissions } (N_0)}$$

Percentage fission yield $Y(A)\% = \frac{N_A}{N_0} \times 100\%$

When this percentage of fission yield is plotted against mass number A for fission chains of ^{235}U, we get a mass distribution of fission fragments (Fig. 27.10).

- The maxima lie near mass numbers 95 and 139 implying that fission is highly asymmetric.
- The yield is minimum for A ~ 117, which corresponds to symmetrical fission (equal masses).
- Some of the fragments may have as low a mass number as 72 or as high as 158.
- The *asymmetric mass distribution* of fission fragments involves the *shell structure* of nuclei and the existence of *magic numbers*. Corresponding to the peak ($A \sim 95$ and 139) yields, the atomic numbers are $Z \sim 42$ and 57. But the Z-value for minimum yield ($A \sim 117$) is $Z \sim 50$ that corresponds to a magic number. Hence its nuclear binding is strong, corresponding to a *closed nuclear shell*. Necessarily, therefore, the fission yield would be minimum. Thus, asymmetric fission is more probable for uranium.

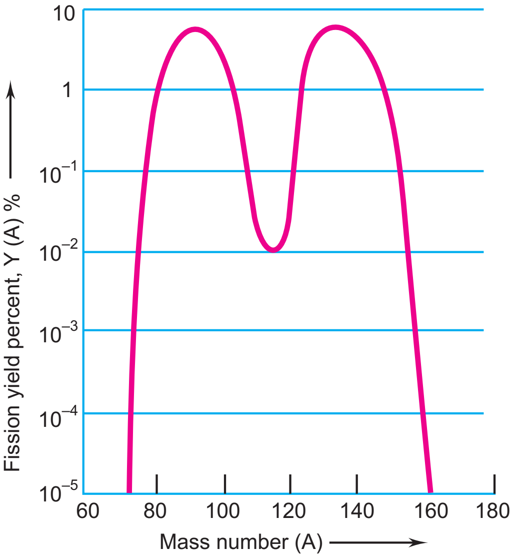


Fig. 27.10

27.15 Applications of Semi-Empirical Mass-Formula

1. Mass Parabolas: Prediction of stability of nuclei against β-decay.

The isobaric nuclei have the same mass number A but different atomic numbver Z.

The semi-empirical mass formula can be written as,

$$ {}_ZM^A = Z(M_p - M_n) + A(M_n - a_v) + a_c\frac{Z^2}{A^{1/3}} + a_s A^{2/3} + a_a A - a_a 4Z + a_a\frac{4Z^2}{A} \pm \delta $$

$$ {}_ZM^A = \alpha A + \beta Z + \gamma Z^2 \pm \delta \qquad ...(1) $$

Here, $\alpha = M_n - \left(a_v - a_a - \frac{a_s}{A^{1/3}}\right)$

$\beta = -4a_a - (M_n - M_p)$

$\gamma = \left(\frac{4a_a}{A} + \frac{a_c}{A^{1/3}}\right)$

δ = 0 for odd-Z, even N

= 0 for even-Z, odd-N

= + δ of odd-Z, odd-N

= – δ for even Z, even-N

± δ is the pairing energy and does not contain terms in Z.

When A is constant, Eq. (1) is the equation of a *parabola*. It is known as the *mass parabola*. Eq. (1) gives the dependence of nuclear mass on the nuclear charge with *constant A*. This dependence is *parabolic*.

The most stable nucleus has the *minimum* mass.

Differentiating Eq. (1) for constant A and equating it to zero, we have

$$\frac{\partial}{\partial Z}\left({}_Z M^A\right) = 0 = \beta + 2\gamma Z_0 \text{ at the point } Z = Z_0.$$

$$\therefore \quad Z_0 = \text{nuclear charge of the ``\textit{most stable}'' isobar} = -\frac{\beta}{2\gamma} \qquad \text{...(2)}$$

(*i*) Mass parabolas of odd *A* isobars

For Odd A nuclei, $\delta = 0$ and thus there is only one parabola, implying that there is only one stable nucleus.

The isobar $Z = Z_0 + 1$ lying on the right hand side of the parabola (Fig. 27.11), has greater mass. It will undergo β^+ decay, giving rise to a sable isobar with $Z = Z_0$.

Similarly, the isobar lying on the left of the curve will undergo β^- decay.

The situation is similar for two isobars $(A, Z_0 + 2)$ and $(A, Z_0 - 2)$ which can decay to $(A, Z_0 + 1)$ and $(A, Z_0 - 1)$, respectively; and so on.

$${}^{73}_{34}\text{Se} \xrightarrow{\beta^+, K} {}^{73}_{33}\text{As} \xrightarrow{K} {}^{73}_{32}\text{Ge} \text{ (stable)}$$

$${}^{73}_{30}\text{Zn} \xrightarrow{\beta^-} {}^{73}_{31}\text{Ga} \xrightarrow{\beta^-} {}^{73}_{32}\text{Ge} \text{ (stable)}$$

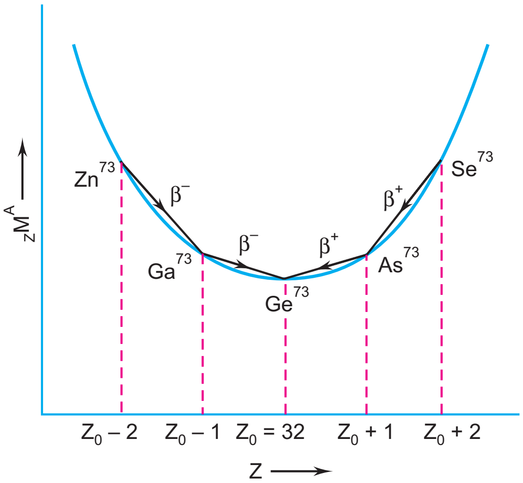


Fig. 27.11

Transition Energies

$$\begin{aligned} {}_{Z_0}M^A &= \alpha A + \beta Z_0 + \gamma Z_0^2 = \alpha A - \gamma Z_0^2 && \text{(using Eq. 2)} \\ {}_ZM^A - {}_{Z_0}M^A &= (\alpha A + \beta Z + \gamma Z^2) - (\alpha A - \gamma Z_0^2) = \gamma (Z - Z_0)^2 && \text{...(3)} \end{aligned}$$

Energy released in $\{Z \to (Z + 1)\}$ transition is

$$Q_{\beta^-} = {}_ZM^A - {}_{Z+1}M^A = \gamma\,[(Z - Z_0)^2 - \{(Z + 1) - Z_0\}^2] = 2\gamma\,(Z_0 - Z - 1/2)$$

The Q value for $Z \to (Z - 1)$ transition is given by

$$Q_{\beta^+} = {}_ZM^A - {}_{Z-1}M^A = \gamma\,\{(Z - Z_0)^2 - (Z - 1 - Z_0)^2\} = 2\gamma\,[Z - Z_0 - 1/2)$$

(*ii*) Mass parabolas of even A Isobars

Here the pairing term $\delta \neq 0$. Since both odd-odd and even-even nuclei are included, we have two parabolas. The lower parabola corresponds to more stable nuclei with even Z, while the upper parabola to less stable nuclei with odd Z. The vertical separation between two parabolas is 2δ due to the opposite sign of δ in two cases.

Fig. 27.12 shows the decay scheme. The reaction energy Q for transition $Z \to Z \pm 1$ for isobars of mass number A is

$$Q_\beta = 2\gamma\,\{\pm[Z_0 - Z] - 1/2\} \pm 2\delta \left| \begin{array}{l} +\,2\delta \text{ for odd } Z \\ -\,2\delta \text{ for even } Z \end{array}\right.$$

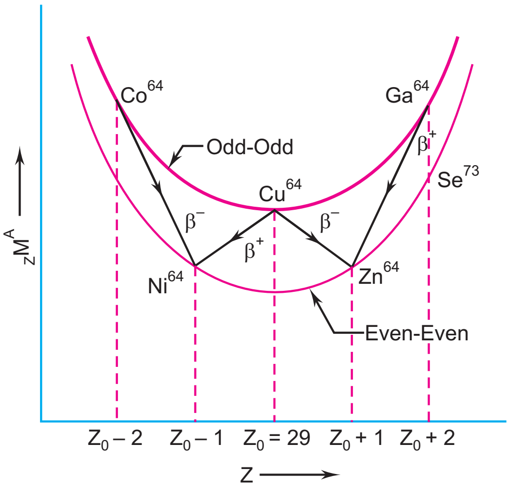


Fig. 27.12

2. Alpha Decay:

The α-decay process is conventionally written as

$${}_ZX^A = {}_{Z-2}Y^{A-4} + {}_2He^4 + Q_\alpha$$

The Q_α-disintegration energy involved in α-decay process is written as

$$Q_\alpha = {}_ZM^A - {}_{Z-2}M^{A-4} - {}_2M^4\,({}_2He^4)$$

In terms of B.E. of nuclei involved, this expression for large Z and A becomes

$$\begin{aligned}
Q_\alpha\,(A, Z) &= \text{B.E.}\,(A-4, Z-2) + \text{B.E.}\,({}_2He^4) - \text{B.E.}\,(A, Z) \\
&= a_v\,(A-4) - a_s\,(A-4)^{2/3} - a_c\,\frac{(Z-2)^2}{(A-4)^{1/.3}} - a_a\,\frac{(A-2Z)^2}{(A-4)} \\
&\qquad - a_vA + a_sA^{2/3} + a_c\,\frac{Z^2}{A^{1/3}} + a_a\,\frac{(A-2Z)^2}{A} + \text{B.E.}\left({}_2He^4\right) \\
&= \text{B.E.}\,({}_2He^4) - 4a_v + a_s\,\{A^{2/3} - (A-4)^{2/3}\} \\
&\qquad + a_c\left\{\frac{Z^2}{A^{1/3}} - \frac{(Z-2)^2}{(A-4)^{1/3}}\right\} + a_a\,(A-2Z)^2\left(\frac{1}{A} - \frac{1}{(A-4)}\right) \\
&= 28.3 - 4a_v + \frac{8}{3}a_sA^{-1/3} + \frac{4a_cZ}{A^{1/3}}\left(1 - \frac{Z}{3A}\right) - \frac{4a_a\,(A-2Z)^2}{A\,(A-4)}
\end{aligned}$$

In the above formulation, pairing energy term has been neglected. If numerical values of a_v, a_s, a_c and a_a are employed, it is found that $Q_\alpha > 0$ for $A > 160$. Therefore, the nuclei $A > 160$ should be α-disintegrating. Practically α-disintegration is observed in nuclei with $A > 200$. For nuclei $A < 200$, the energy release is so small that barrier penetration becomes small.

EXAMPLE 3. *What are mass parabolas? Define mirror nuclei and obtain an expression for* β *disintegration energy of mirror nuclei.* (*Mumbai University, October 2010*)

Beta Disintegration Energy of Mirror Nuclei:

Mirror nuclei are pairs of isobaric nuclei in which the proton and neutron numbers are interchanged and differ by one unit. ${}_6C^{11}$ and ${}_5B^{11}$ are mirror nuclei. For such pairs $A = 2Z - 1$.

The members of the pairs with higher Z are usually found β^+ emitters such as

$${}_6C^{11} \longrightarrow {}_5B^{11} + \beta^+ + \nu$$

The maximum energy of the β^+ emitted in this case is 0.96 MeV. No γ-ray energy is emitted in the transition.

We neglect pairing energy term δ (for odd A).

The semi-empirical mass formula is written as

$$\begin{aligned}
M\,(A, Z) &= Z\,M_p + NM_n - a_vA + a_sA^{2/3} + a_c\,\frac{Z^2}{A^{1/3}} + a_a\,\frac{(A-2Z)^2}{A} \\
&= Z\,M_p + (Z-1)\,M_n - a_vA + a_sA^{2/3} + a_c\,\frac{Z^2}{A^{1/3}} + \frac{a_a}{A}
\end{aligned}$$

For the daughter term

$$M\,(A, Z-1) = (Z-1)\,M_p + (Z-2)\,M_n - a_vA + a_sA^{2/3} + a_c\,\frac{(Z-1)^2}{A^{1/3}} + \frac{a_a}{A}$$

$$M\,(A, Z) - M\,(A, Z-1) = M_p - M_n + \frac{a_c\,(2Z-1)}{A^{1/3}}$$

$$\begin{aligned}\therefore \qquad Q_{\beta^+} &= M(A, Z) - M(A, Z-1) - 2m_e \\ &= M_p - M_n + \frac{a_c\,(2Z-1)}{A^{1/3}} - 2m_e \\ &= a_c A^{2/3} - (M_n - M_p + 2m_e) \\ Q_{\beta^+} &= a_c A^{2/3} + 1.804 \text{ MeV}\end{aligned}$$

The equation shows that plot of disintegration energy Q_{β^+} against $A^{2/3}$ should be straight line with a slope $a_c = \dfrac{3e^2}{20\pi\varepsilon_0 r_0}$. The slope gives nuclear radius of the order 1.44 fermi.

EXERCISE

1. Find the energy release, if two ${}_1\text{H}^2$ nuclei can fuse together to form ${}_2\text{He}^4$ nucleus. The binding energy per nucleon of H^2 and He^4 is 1.1 MeV and 7.0 MeV respectively.

[**Hint.** B.E. for ${}_2\text{He}^4$ = 28 MeV; B.E. for ${}_1\text{H}^2$ nucleus = 2.2 MeV.

$$\begin{aligned}\Delta E &= [\text{Mass of } {}_1\text{H}^2] - [\text{Mass of } {}_2\text{He}^4] \\ &= 2\,[M_p + M_n - 2.2] - [2M_p + 2M_n - 28.0] = 23.6 \text{ MeV}]\end{aligned}$$

2. Calculate the binding energy per nucleon in ${}_6\text{C}^{12}$. Masses of proton, neutron and electron are 1.007276, 1.008665 and 0.00055 u respectively. The mass of ${}_6\text{C}^{12}$ atom is 12.000000 u.

[**Hint.** mass of 6 protons + mass of 6 neutrons = (6 × 1.007276 + 6 × 1.008665) u = 12.095646 u.

$$\begin{aligned}\text{Mass of carbon nucleus} &= \text{mass of carbon atom} - \text{mass of 6 electrons} \\ &= (12.000000 - 6 \times 0.00055)\,u = 11.996700\,u \\ \Delta m &= (12.095646 - 11.996700)\,u = 0.098946\,u \\ \text{B.E.} &= 0.098946 \times 931.3 \text{ MeV} = 92.148 \text{ MeV} \\ \text{B.E. per nucleon} &= 92.148/12 = 7.69 \text{ MeV}]\end{aligned}$$

3. The deviation of the charge distribution of a nucleus from spherical symmetry can be estimated by measuring its

(*a*) electric charge
(*b*) electric dipole moment
(*c*) magnetic dipole moment
(*d*) electric quadrupole moment.

[**Ans.** (*d*)]

4. *Give an account of the methods of separating isotopes.*

Ans. Since isotopes have the chemical properties, they cannot be separated from one another by chemical methods. Isotopes can be separated from one another only by making use of the differences in their physical properties. All the methods of separation are based on the difference in the masses of the isotopes.

(1) *Mass spectrograph method*. The photographic plate in the mass spectrograph of Aston's type is replaced by different vessels where the different isotopes can be separately collected. Complete separation is possible by this method, but the yield is extremely small. The main difficulty is that extremely high currents of ions or long intervals of time are required to separate any appreciable amount of an element into its isotopes.

(2) *Diffusion method*. The rate of diffusion of a gas through a porous partition is inversely proportional to the square root of the masses of the molecules. Hence different isotopes of a gas will diffuse with slightly different velocities.

Fig. 27.13 shows Hertz multiple-stage porous-tube apparatus for separating isotopes by diffusion.

The progress of the main volume of gas is from right to left, as it becomes enriched in the way heavy fractions. The lighter fractions diffuse out through the porous tubes (crosshatched) and are returned by the recirculating pumps P to the previous stage, eventually collecting in the reservoir L. The heavier fractions progress from state to stage, eventually collecting in the reservoir H. The over-all enrichment increases exponentially with the number of stages. Depending on the isotopes to be separated, some 10 to 50 or more stages may be used.

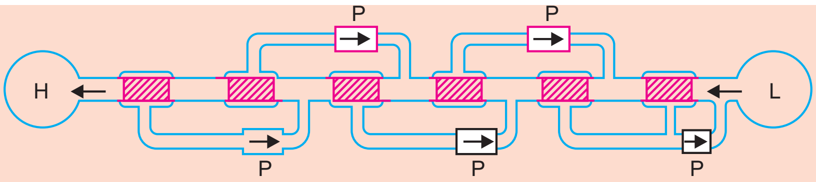


Fig. 27.13

(3) *Thermal diffusion method*. This is based on the fact that if a gaseous mixture of isotopes is placed in a vessel, one part of which is hotter than the other, the lighter molecules collect in the region of higher temperature.

Fig. 27.14 show the Clusius-Dickel thermal-diffusion isotope separator. A cooled vertical glass tube with an electrically heated wire along its axis is used. By thermal diffusion the heavier fraction tends to concentrate at the cool outer wall.

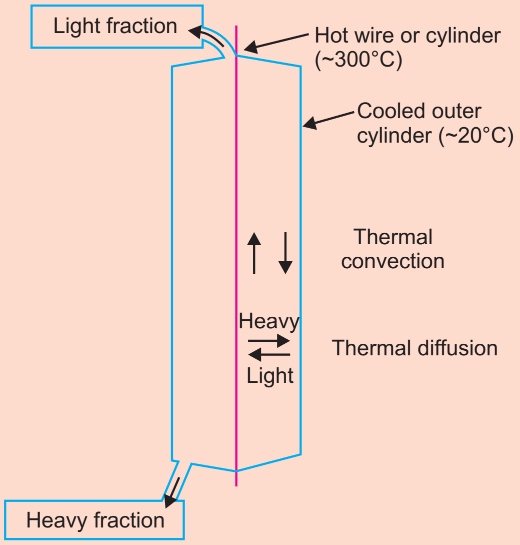


Fig. 27.14

The lighter fraction concentrates at the hot inner cylinder or wire. The action of gravity then causes thermal convection which provides an effective downward transport for the heavy fraction at the cool outer wall and an upward transport for the light fraction along the axis. The pressure is maintained at a low enough value to avoid turbulence in the thermal convective flow.

The process separation factor α, in terms of the equilibrium mole concentrations in the heavy and light reservoirs, is

$$\alpha = e^{lA(M_2 - M_1)/(M_2 + M_1)}$$

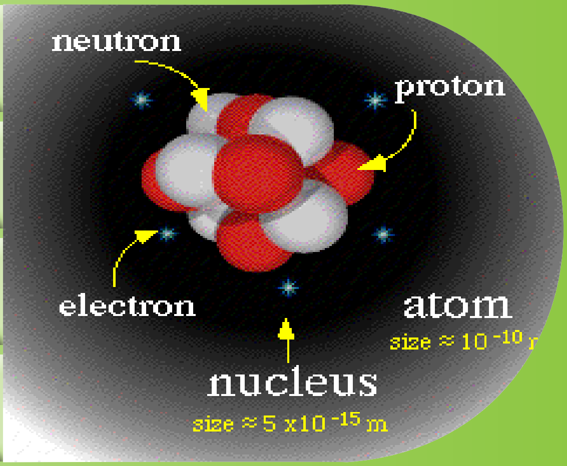
Here, M_1 and M_2 are the masses of the light and heavy isotopic molecules being separated, l is the length of the column.

A is a function of the viscosity, self-diffusion, and density of the gas, the temperatures and radii of the cylindrical walls, and the gravitational constant.

(4) *Pressure diffusion (or centrifugal) method*. This is based on the fact that if a gas or vapour flows into a rapidly rotating cylinder, the force acting on the molecules will result in an increased concentration of the heavier isotopes at the walls while the lighter ones collect near the axis of rotation. Isotopes of chlorine in carbon tetrachloride and of bromine in ethyl bromide have been separated by this method.

CHAPTER

28


neutron
proton
electron
atom
size ≈ 10 -10 r
nucleus
size ≈ 5 x10 -15 m

NUCLEAR STRUCTURE

AT A GLANCE

28.1	Measurement of Nuclear Radius	**28.2**	Parity
28.3	Statistics of Nuclei	**28.4**	Properties of Ground State of Deuterium Nucleus
28.5	Nuclear Models—The Shell Model	**28.6**	Fermi Gas Model of the Nucleus

28.1 Measurement of Nuclear Radius

Fig. 28.1 shows Hofstadter's experimental arrangement for nuclear radius measurement. This method consists of measurement of nuclear charge distribution and assumes that charge distribution and matter distribution within nuclei are essentially the same. The charge distribution in nuclei is obtained by analysing scattering of high-energy electrons from various nuclei. The elastic-scattering process is

$$e + X \rightarrow e + X. \qquad \text{...(1)}$$

Electron is chosen as the bombarding particle because its interaction with the nucleus, the familiar electro-magnetic interaction, is so well-known.

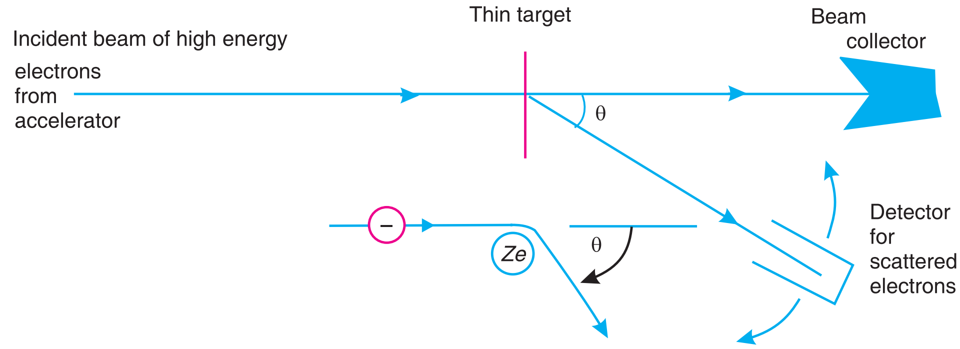
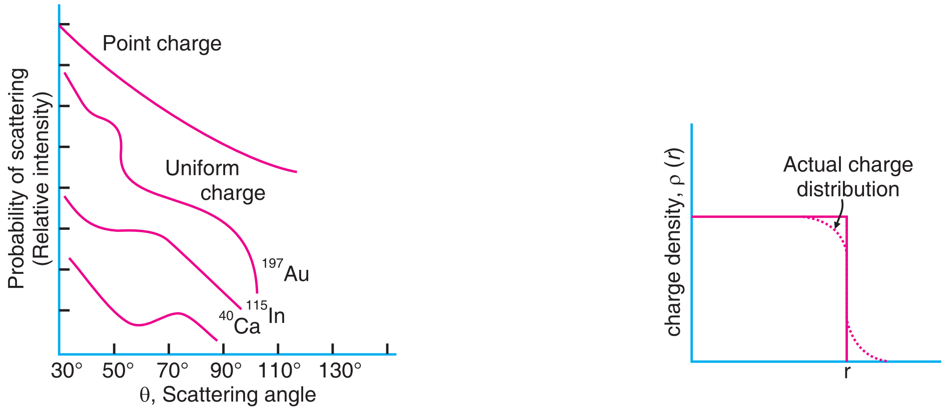

Thin target
Beam collector
Incident beam of high energy electrons from accelerator
θ
–
Ze
θ
Detector for scattered electrons

Fig. 28.1

The equipment includes an electron accelerator to prepare the high energy electron beam, a scattering target of species X, and a spectrometer to detect electrons scattered elastically in directions given by the indicated scattering angle θ.



(*a*) (*b*)

Fig. 28.2

The de Broglie wavelength of 200 MeV electrons is about 10^{-14} m which is the size of a nucleus. Thus to get information about nuclear charge distribution, the incident electron beam energy has to be of the order of 200 MeV. The assumption that nuclear charge is uniformly spread over a spherical volume and not a point charge leads one to expect diffraction effects. Portions of electron waves incident on different parts of the nucleus will be scattered in a particular direction, with phase differences resulting in constructive or destructive interference at some angle. Figure 28.2 (*a*) shows the angular distribution of elastically scattered 200 MeV electrons from nuclei assumed to have spherical uniform charge distribution (having uniform density up to radius r). The actual nuclear charge distribution is not exactly given by a step function, but is shown by the dotted line [Fig. 28.2 (*b*)].

Hofstadter assumed that the nuclear charge density has the form

$$\rho(r) = \frac{\rho_0}{1 + e^{(r-R)/b}} \qquad \text{...(2)}$$

Here, ρ_0 is density at the nuclear centre, R is the radius at which ρ falls to $\rho_0/2$ and b measures how rapidly ρ falls to zero at the nuclear surface.

Nuclear charge distribution found in some nuclei using high-energy electrons as probes is shown in Fig. 28.3. This analysis determines a nuclear radius R that grows with mass number A according to the formula

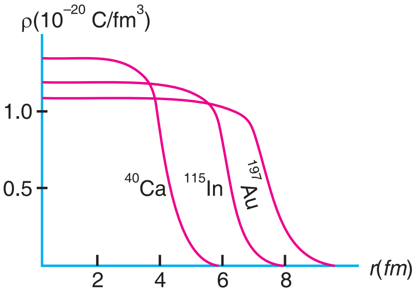


Fig. 28.3

$$R = R_0 A^{1/3} \qquad \text{...(3)}$$

28.2 Parity

The total spin of a nucleus consists of the sum of orbital angular momentum of nucleons and the sum of their spins. The orbital angular momentum $\sum_n l_n$ actually defines the parity of the nuclei. The angular momentum eigen function can be expressed as a Spherical Harmonics Y_m which is an even function for even l_n and an odd function for an odd l_n. This shows that $\sum_n l_n$ is either even or odd for nuclei. According to this, the nuclear wave functions are said to have either an even parity or odd parity.

For positive or even parity, $\psi(x, y, z) = \psi(-x, -y, -z)$.

For negative or odd parity, $\psi(x, y, z) = -\psi(-x, -y, -z)$.

If $\sum_n l_n$ is even, the parity of nuclide is positive and if $\sum_n l_n$ is odd, the parity is negative. As an example, the Deuteron nucleus contains a neutron and a proton in S-state with $l = 0$ and the parity of deuteron is positive.

Nuclei of various atoms in a ground state have a definite parity which is either positive or negative. When the nuclei are in an excited state, their parities are not always the same as in the ground state. Parity in nuclear transformations is conservable quantity but is not conserved in weak interactions like Beta decay. It is conserved in nuclear reactions and gamma decay.

28.3 Statistics of Nuclei

Since nuclear particles like protons are identical, the probability distribution $|\psi|^2$ remains unchanged by an exchange of particle co-ordinates. If x_1 are the general co-ordinates of one particle (which may include spin) and x_2 are the general co-ordinates for the other, then we have,

$$|\psi(x_1, x_2)|^2 = |\psi(x_2, x_1)|^2$$

$$\psi(x_1, x_2) = \pm\psi(x_2, x_1).$$

The sign we choose depends on the type of particle we are considering. For electrons, protons, etc., we must choose the negative sign, and the eigen function for the two electrons is antisymmetric for an exchange of coordinates. Such particles are said to satisfy *Fermi-Dirac statistics* and are called *fermions.* On the other hand, for photons, one must choose the positive sign so that the eigen function of two photons is symmetric for an exchange of coordinates. Photons are said to satisfy *Bose-Einstein statistics* and are called *bosons.*

In systems obeying Fermi-Dirac statistics, according to Pauli, not more than one particle can occupy the same quantum state. If spin also is taken as a coordinate in describing a particle, two similar particles with spins oriented in opposite way can occupy a quantum state. Neutrons and protons (nucleons) are spin $\frac{1}{2}$ particles and obey Fermi-Dirac statistics. The wave function of a nucleus with A particles will be odd or even according as $(-1)^A$ is positive or negative. Nuclei with odd A will obey Fermi-Dirac statistics and with even A Bose-Einstein statistics under this criterion.

28.4 Properties of Ground State of Deuterium Nucleus

(*i*) Binding Energy. Deuterium nucleus is the simplest among the atomic nuclei. It contains one proton and one neutron. These two particles are bound together with a binding energy of 2.225 MeV. There are essentially two direct methods of experimental measurement of the binding energy of deuteron. One is to measure the energy of gamma radiation emitted when a neutron and proton combine to form a nucleus of deuteron ($n - p$ capture reaction). The second method consists of measuring the energy of gamma radiation that will break the bond between the neutron and proton (photo disintegration of deuteron).

(*ii*) Angular Momentum and Parity. The angular momentum quantum number of the ground state of the deuteron is $\mathbf{I} = 1$. This is obtained as the resultant of the spin of proton (1/2) and of neutron (1/2). The orbital angular momentum for both the nucleons will be zero. The parity is taken as + ve, which is corroborated by experimental results.

***(iii)* Magnetic Dipole Moment.** The magnetic dipole moment of the deuteron is + 0.85739 μ_N as given by N.M.R. data. The magnetic dipole moment of a proton is + 2.79275 μ_N and that of a neutron is – 1.91315 μ_N. The magnetic dipole moment of deuteron is slightly smaller than the resultant obtained by the algebraic sum of the magnetic moments of proton and of neutron.

***(iv)* Electric Quadrupole Moment.** The electric quadrupole moment of the deuteron is Q = + 0.00282 barn. The quantity of the quadrupole moment is an indication that the charge distribution is not spherically symmetrical in deuteron. The + ve sign for Q indicates that the charge distribution is prolate rather than oblate.

***(v)* Radius of Deuteron.** The size of the deuteron is found by studying the scattering of high energy electrons. The average radius of the deuteron is found to be 4.2 fermi. [It is the root mean square distance between the neturon and proton].

***(vi)* The Force between a Neutron and Proton is Spin-dependent.** I, Angular momentum of deuteron = 1 in units of $\hbar$. *Since the spins of both, the neutron and the proton, is* $\frac{1}{2}$ *in the units of* $\hbar$*, the deuteron must be formed with the spin vectors of the neutron and the proton parallel. The* other possibility of the two spin vectors antiparallel giving a net spin = 0 *does not lead to a bound state for the deuteron.* Deuteron excited states do not exist and so the state to be studied is the *ground state*. For the ground state, $l = 0$. It is 1*s* state.

$$\therefore \qquad I = L + S = S = 1.$$

The 1*s* state (ground state) with $S = 0$ does not give a bound state for deuteron. The energy of the ground state of the deuteron is different if the spins of the neutron and proton are parallel than if they are anti-parallel. Otherwise why should the spin parallel system be stable and spin anti-parallel not stable? The important conclusion that can be drawn from this is that *the force between a neutron and a proton is spin-dependent. The* nuclear force keeping the two nucleons system bound, thus depends on whether the spins of the two nucleons are parallel or antiparallel *i.e.*, whether the total spin $S = 1$ or 0. This feature is quite different than, for example, the Coulomb force which goes as $1/r^2$ and so is just a function of r.

28.5 Nuclear Models—The Shell Model

Nuclear Energy Level Scheme and Explanation of Magic Numbers. To account for the observed magic numbers, Mayer and Jensen postulated a strong *nuclear spin-orbit interaction*. The magnitude of the spin-orbit interaction is such that the consequent splitting of energy levels into sublevels is many times larger than the analogous splitting of atomic energy levels. The nuclear spin-orbit splitting of a single-nucleon energy level is assumed to be *large* and also *inverted* (Fig. 28.4). We ascribe this behaviour to a nuclear interaction of the form

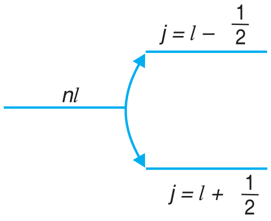


Fig. 28.4

$$V_{SL}\,(\text{nucleus}) = -\frac{a_{SL}^2}{r}\frac{dV}{dr}\mathbf{S.L}.$$

The minus sign accomplishes the required inversion of the split levels.

The constant a_{SL}^2 produces the desired amount of energy splitting. The central-field function $V(r)$ appears along with the orbital and spin angular momenta of the nucleon. The exact form of the potential-energy function is not critical, provided that it more or less resembles a square well.

The shell theory assumes that *LS* coupling holds only for the very lightest nuclei, in which the l values are necessarily small in their normal configurations. In this scheme, the intrinsic spin angular momenta S_i of the particles concerned are coupled together into a total spin momentum **S**. The orbital angular momenta $\mathbf{L}_i$ are separately coupled together into a total orbital momentum **L**. Then **S** and **L** are coupled to form a total angular momentum **J** of magnitude $\sqrt{\mathbf{J}\,(\mathbf{J}+1)}\,\hbar$.

After a transition region in which an intermediate coupling scheme holds, the heavier nuclei exhibit **jj coupling**. In this case, the $\mathbf{S}_i$ and $\mathbf{L}_i$ of each particle are first coupled to form a $\mathbf{J}_i$ for that particle of magnitude $\sqrt{\mathbf{J}(\mathbf{J}+1)}\,\hbar$. The various $\mathbf{J}_i$ then couple together to form the total angular momentum **J**. The *jj* coupling scheme holds for the great majority of nuclei.

Fig. 28.5 shows the nucleon energy levels according to the shell model. The levels are designated by a prefix equal to the total quantum number *n*, a letter that indicates *l* for each particle in that level, and a subscript equal to *j*. The spin-orbit interaction splits each state of given *j* into $2j+1$ substates. The accumulated population of nucleons corresponds to a magic number at every one of the larger energy gaps. Hence shells are filled when there are 2, 8, 20, 28, 50, 82 and 126 neutrons or protons in a nucleus.

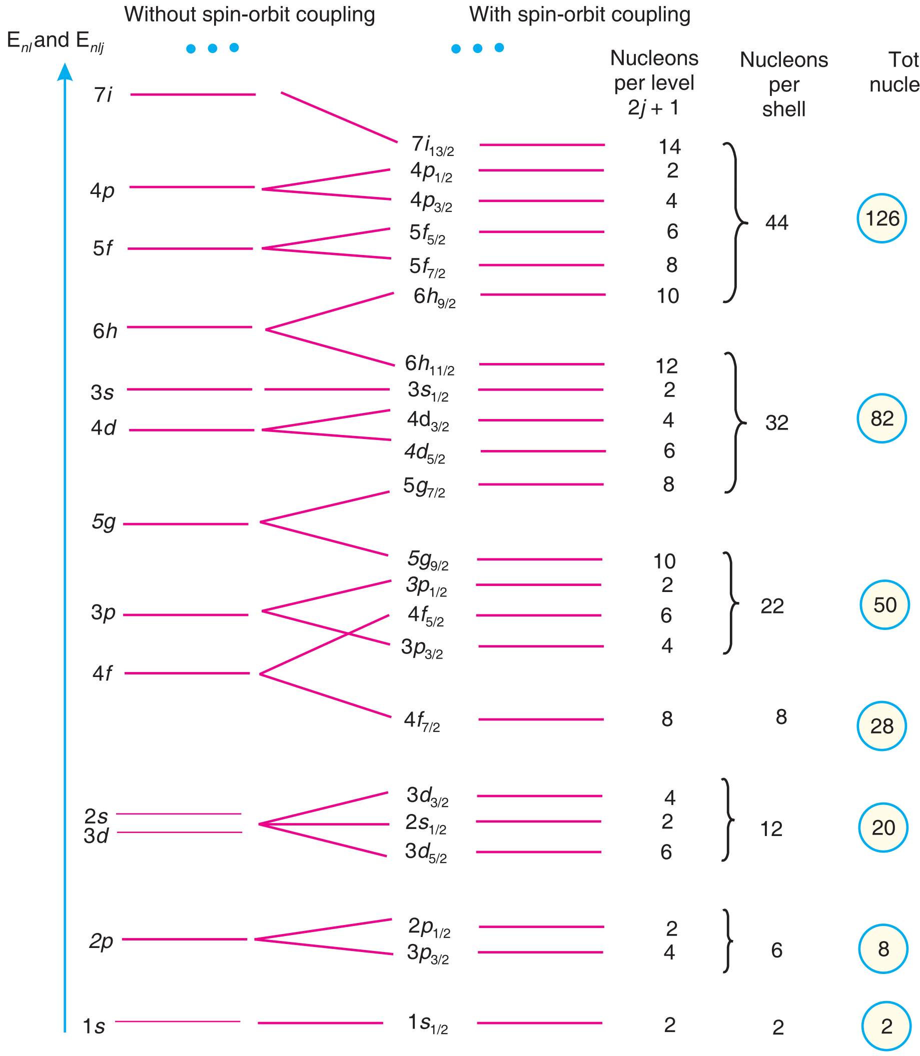

Without spin-orbit coupling
With spin-orbit coupling
E_{nl} and E_{nlj}
Nucleons per level $2j+1$
Nucleons per shell
Total nucleons
7i
4p
5f
6h
3s
4d
5g
3p
4f
2s
3d
2p
1s
$7i_{13/2}$ 14
$4p_{1/2}$ 2
$4p_{3/2}$ 4
$5f_{5/2}$ 6
$5f_{7/2}$ 8
$6h_{9/2}$ 10
44
126
$6h_{11/2}$ 12
$3s_{1/2}$ 2
$4d_{3/2}$ 4
$4d_{5/2}$ 6
$5g_{7/2}$ 8
32
82
$5g_{9/2}$ 10
$3p_{1/2}$ 2
$4f_{5/2}$ 6
$3p_{3/2}$ 4
22
50
$4f_{7/2}$ 8
8
28
$3d_{3/2}$ 4
$2s_{1/2}$ 2
$3d_{5/2}$ 6
12
20
$2p_{1/2}$ 2
$3p_{3/2}$ 4
6
8
$1s_{1/2}$ 2
2
2

Fig. 28.5

28.6 Fermi Gas Model of the Nucleus

The degenerate gas model was suggested by Fermi and pictures the nucleus as a gas of neutrons and protons. In this model, the forces between pairs of nucleons and also the surface effects like capillarity are omitted. Protons and neutrons exist in a box of nuclear dimensions and fill the lowest available quantum states to the extent permitted by the exclusion principle. Since both protons and neutrons have spins of 1/2, they are fermions and obey Fermi-Dirac statistics.

It is assumed that the nucleons are confined to a small volume equal to $\frac{4}{3}\pi R_0^3 A$. Under these circumstances the nucleus is completely degenerate even in the first few excited states *i.e.*, unlike a classical gas, it occupies almost all the lowest energy states available. This is so because the spacing of energy levels in a gas confined to such a small volume as that of nuclear dimensions is of the order of several MeV. As the excitation energies of the first few excited states are not greater than this, the ground state and the first few excited states are therefore completely degenerate.

The de Broglie wave-length for a nucleus of radius R and A particles of mass M is, $\lambda = R/A^{1/3}$. The corresponding momentum

$$p = \frac{h}{\lambda} = \frac{2\pi\hbar}{\lambda} = \frac{2\pi\hbar A^{1/3}}{R} = 2\pi\hbar(A^{1/3}/R).$$

The kinetic energy per particle will be proportional to $\frac{p^2}{2M}$.

$$\text{K.E} \propto \frac{A^{2/3}}{M R^2}.$$

For the whole nucleus, $\text{K.E} \propto \frac{A^{5/3}}{M R^2}$.

The potential energy is proportional to the number of interacting pairs *i.e.*, to $A\,(A-1)/2$.

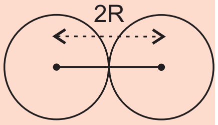
For A large enough, the potential energy is the main term. The nucleus would therefore collapse according to this model. The properties of neutron-proton interaction must be postulated to prevent this.

This model gives a good qualitative picture of the nucleus specially of heavy nuclei but the actual numerical results *e.g.*, energy levels are inaccurate.

EXERCISE

1. Mention various methods for determining the size of the nucleus and describe any one in detail. (*Meerut* 1980, 86)
2. Explain the terms, (*a*) Parity and (*b*) Statistics of nuclei.
3. Calculate the parity of deuteron, Li^7, B^{10} and N^{15} nuclei.
4. Explain the basic properties of deuteron nucleus.
5. Explain with arguments how the force between a proton and a neutron is spin dependent.
6. Explain how the shell model of a nucleus accounts for the existence of magic numbers.
7. Write a short-note on the Fermi gas model of the nucleus.

8. Calculate the electrostatic potential energy between two equal nuclei produced by the fission of ${}_{92}U^{235}$ at the moment of separation.



Sol. Before fission, ${}_{92}U^{235} + {}_0n^1$. *i.e.*, A = 235 + 1 = 236 and Z = 92.

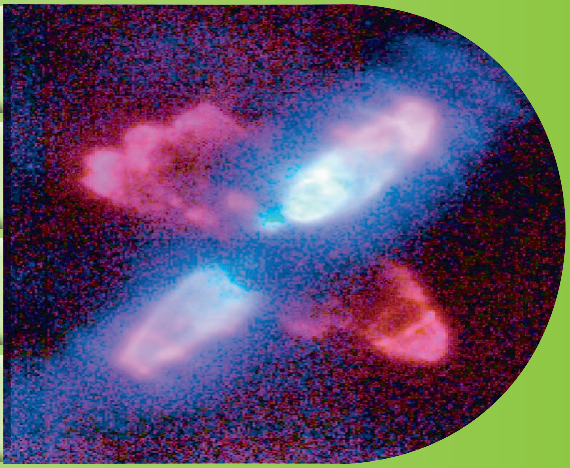
Two equal nuclei produced will each have A = 118 and Z = 46.

The radius of the nucleus, $R = R_0A^{1/3} = (1.3 \times 10^{-15})\ (118)^{1/3} = 6.376 \times 10^{-15}$ m

The electrostatic potential energy between the product nuclei is given by

$$E = \frac{1}{4\pi\epsilon_0}\frac{(Ze)(Ze)}{(2R)} = \frac{(46)^2(1.6\times10^{-19})^2}{4\pi\times(8.854\times10^{-12})\times2\times(6.376\times10^{-15})}$$

$$= 3.81952\times10^{-11}\,J = \frac{3.81952\times10^{-11}}{1.6\times10^{-13}}\,MeV = \mathbf{238.72\ MeV}$$



CHAPTER
29

DETECTORS OF NUCLEAR RADIATIONS

AT A GLANCE

29.1	Introduction		
29.2	Interaction between Energetic Particles and Matter		
29.3	Ionization Chamber	29.4	Solid-state Detectors
29.5	Proportional Counter	29.6	Geiger-Muller Counter
29.7	The Wilson Cloud Chamber	29.8	Diffusion Cloud Chamber
29.9	Bubble Chamber	29.10	Spark Chamber
29.11	Nuclear Emulsions	29.12	The Scintillation Counters
29.13	Cerenkov Counter		

29.1 Introduction

Most of the nuclear reactions are accompanied by the emission of charged particles like α- particles, protons, electrons and radiations like γ-rays. In order to understand these particles and their interaction with atomic nuclei, precise information about their mass, momentum, energy, etc., are necessary. We shall describe in this chapter some of the common techniques employed for the detection of nuclear radiations and for analysing their energies.

Several nuclear radiation detectors depend for their operation on the *ionization* that is produced in them by the passage of charged particles. This group of detectors includes ionization chambers, proportional counters, G-M counters, semiconductor radiation detectors, cloud chambers and spark chambers. In other detectors, excitation and sometimes molecular dissociation also play important roles. These phenomena, in combination with ionization, bring about the luminescence in scintillation detectors and the latent images in photographic emulsions.

29.2 Interaction between Energetic Particles and Matter

(*a*) Heavy Charged Particles. A heavy charged particle (like a proton, α-particle or fission fragment) has a fairly definite *range* in a gas liquid, or solid. The particle loses energy primarily by the excitation and ionization of atoms in its path. The energy loss occurs in a large number of small increments. The primary particle has such a large momentum that its direction is usually not seriously changed during the slowing process. Eventually it loses all its energy and comes to rest. The distance traversed is called the *range* of the particle.

The energy loss per unit length ($-dE/dx$) is called the *stopping power.* The rate $-dE/dx$ at which a heavy particle of charge ze and speed v loses energy in an absorber of atomic number Z which contains N atoms per unit volume whose average ionization energy is I is given by

$$-\frac{dE}{dx} = \frac{z^2 e^4 N Z}{4\pi\varepsilon_0{}^2 m_0 v^2}\left[\text{In}\left(\frac{2m_0 v^2}{I}\right) - \text{In}\left(1-\frac{v^2}{c^2}\right) - \frac{v^2}{c^2}\right] \qquad \text{...(1)}$$

m_0 is the electron rest mass.

The range can be calculated by integrating Eq. (1) over the energies of the particle

$$R = \int_T^0 \left(-\frac{dE}{dx}\right)^{-1} dE \qquad \text{...(2)}$$

(*b*) Electrons. Electrons interact through coulomb scattering from atomic electrons, just like heavy charged particles. There are however, a number of important differences:

(1) Electrons travel at relativistic speeds.

(2) Electrons will suffer large deflections in collisions with other electrons, and therefore will follow erratic paths. The range will therefore be very different from the length of the path that the electron follows.

(3) Very energetic electrons (E > 1 MeV) lose an appreciable fraction of their energies by producing continuous X-rays (also called *Bremsstrahlung*). The cross section for this process increases with increasing E.

(*c*) The Absorption of γ-Rays. The interaction of γ-rays with matter is markedly different from that of charged particles such as α or β particles. γ-rays are extremely penetrating so that they are able to pass through considerable thicknesses of matter. γ-rays show the exponential absorption in matter. If radiation of intensity I is incident upon an absorbing layer of thickness dx, the amount of radiation absorbed dI is proportional both to dx and I. Hence,

$$dI = -\mu I\, dx \text{ or } I = I_0\, e^{-\mu x}.$$

Here, μ is a constant of proportionality which is a characteristic property of the medium, known as *linear absorption coefficient.* The *mass absorption coefficient* μ_m may be obtained by dividing μ by the density of the medium. $\mu_m = \mu/\rho$. The above relation gives the intensity (number of quanta per unit area per second) of the beam of initial intensity I_0, after traversing a thickness x of the homogeneous material. At low energies (0.1 MeV to 25 MeV) there are three important processes through which γ photons are absorbed by matter. [Fig. 29.1].

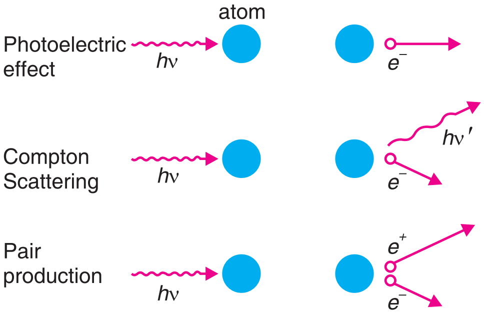


Fig. 29.1

(*i*) Photoelectric effect. In this process, the γ-rays knock out electrons from inside the atoms of the absorbing material. This results in the ionisation of the atoms and the emission of fluorescent radiations. Einstein's equation for those photo-electrons will be

$$\frac{1}{2}mv_k^2 = h\nu - W_k, \frac{1}{2}mv_L^2 = h\nu - W_L \text{ etc.}$$

where $h\nu$ is the photon energy, v_k, v_L represent the velocities of the photoelectrons arising in the K, L... shells, and W_k, W_L... are the binding energies of K, L... shells.

(*ii*) Pair Production. In this process, the photon disappears and is converted to an electron-positron pair. This process can take place only when the photon energy exceeds $2m_0c^2$. The pair production process cannot occur in free space and usually takes place in the presence of a nuclear field. The nucleus recoils in this process conserving momentum. But the K.E., carried away by the nucleus is negligibly small due to its large mass compared with that of the electron. Photon energy, if any, in excess of $2m_0c^2$ is shared as K.E. by the product particles.

(*iii*) Compton effect. It is an elastic scattering process in which the photon imparts energy to an electron. When a photon of energy $h\nu$ strikes the perfectly free electron (at rest), the photon with diminished energy $h\nu'$ is scattered at an angle θ with the direction of incident photon and the electron recoils at an angle ϕ. The energy absorbed by these Compton electrons is only a small fraction of the total energy of the incident γ-rays, unlike in the case of photoelectrons.

At low photon energies, the photoelectric effect is the chief mechanism of energy loss. The importance of the photoelectric effect decreases with increasing energy. In the lighter elements, Compton scattering becomes dominant at photon energies of a few tens of keV, whereas in the heavier ones this does not happen until photon energies of nearly 1 MeV are reached. Pair production becomes increasingly likely the more the photon energy exceeds the threshold of 1.02 MeV. Fig. 29.2 is a graph of the linear attenuation coefficient for photons in lead as a function of photon energy. The contributions to μ of the photoelectric effect, Compton scattering, and pair production are shown.

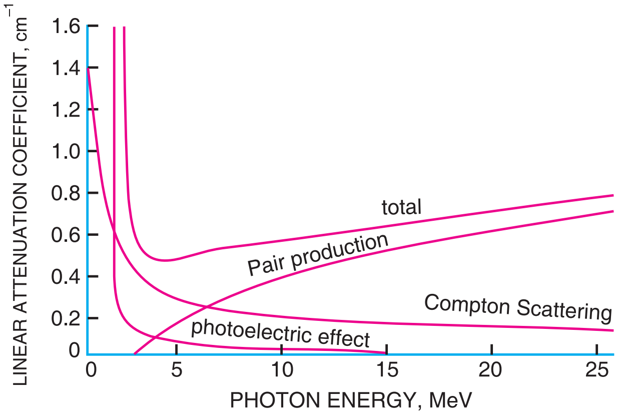


Fig. 29.2

EXAMPLE. *The linear attenuation coefficient for 2-MeV gamma rays in water is about* $5m^{-1}$. (*a*) *Find the relative intensity of a beam of* 2 *MeV gamma rays after it has passed through* 0.1 *m of water.* (*b*) *How far must such a beam travel in water before its intensity is reduced to* 1 *per cent of its original value*?

SOL. We have, $I = I_0 e^{-\mu x}$.

(*a*) Here, $\mu = 5m^{-1}$; $x = 0.1$ m; $I/I_0 = ?$

$$\frac{I}{I_0} = e^{-\mu x} = e^{-0.5} = 0.61.$$

(*b*) Here, $I_0/I = 100$, $\mu = 5m^{-1}$; $x = ?$

$$x = \frac{\log_e(I_0/I)}{\mu} = \frac{\log_e 100}{5} = 0.92\,m.$$

29.3 Ionization Chamber

Ionization Chamber.

The principle employed here is that charged sub-atomic particles can ionise gases. The number of ion-pairs produced gives us information not only on the nature of the incident particles, but even on their energy. The ionisation chamber consists of a hollow metallic cylinder *C*, closed at both ends, with a window *W* at an end for the entry of the ionising particles or radiation [Fig. 29.3]. A metal rod *R*, well insulated from the cylinder, is mounted coaxially within the cylinder. *R* is connected to a quadrant electrometer *E*. A p.d. of several hundred volts is maintained between *C* and *R*. An earthed guard ring *G* prevents leakage of charge from the cylinder to the rod. The chamber contains some gas like sulphur dioxide or methyl bromide. When a charged particle enters the chamber, it produces a large number of ion pairs in the enclosed gas, along its path. Positive ions move towards *R* and negative ions towards *C*.

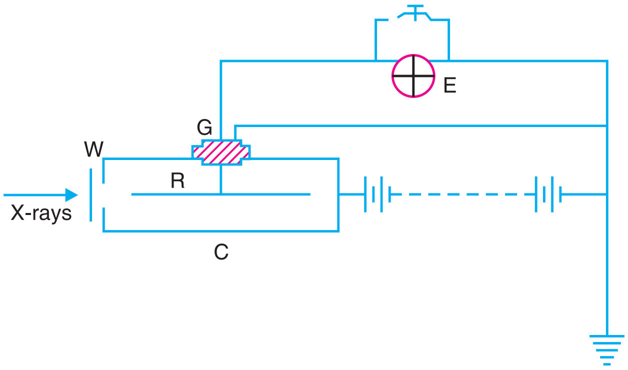


Fig. 29.3

The quadrant electrometer *E* measures the rate of deposition of positive charges on *R*. The ionisation currents produced are quite small $\approx 10^{-12} - 10^{-15}$ amperes. Special electrometers and D.C. amplifying devices have to be employed to measure such small currents.

If individual particles are to be counted, then the pulses of current produced are fed to a pulse amplifier, which is joined to the ionisation chamber by a coupling capacitor [Fig. 29.4]. Ionisation chambers have been used to study α-particles, β-particles, protons, electrons and nuclei of lighter elements. Ionisation chambers were extensively used in the early studies of cosmic ray phenomena. Ionisation chambers can also be used for measurements on X-rays and γ-rays. For neutron detection, the chamber is filled with boron trifluoride vapour (where the boron is enriched with B^{10}) or the chamber walls are lined inside with a boron compound in the form of a paste.

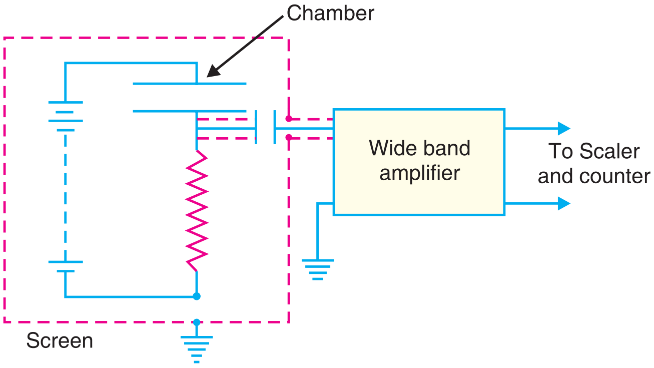


Fig. 29.4

An ionisation chamber is much less sensitive to β-particles (in comparison to α-particles) because β-particles produce fewer pairs of ions in their passage through the chamber.

For detecting γ-rays, an ionisation chamber of thick wall made of high atomic-number material (Pt, Bi) is employed. The γ-rays impinging on the *walls* of the chamber eject high-speed electrons which produce ionisation in the gas.

Example 1. *α-particles of energy* 5 *MeV pass through an ionisation chamber at the rate of* 10 *per second. Assuming all the energy is used in producing ion pairs, calculate the current produced.* (35 *eV is required for producing an ion pair and* $e = 1.6 \times 10^{-19}$ *C*).

Sol. Energy of α-particles = 5×10^6 *eV*.

Energy required for producing one ion pair = 35 *eV*

No. of ion pairs produced by one α-particle

$$= \frac{5\times10^6}{35} = 1.429\times10^5$$

Since 10 particles enter the chamber in one second,

No. of ion pairs produced per second

$$= 1.429 \times 10^5 \times 10 = 1.429 \times 10^6.$$

$$\text{Charge on each ion} = 1.6 \times 10^{-19}\ C.$$

$$\therefore \quad \text{Current} = (1.429 \times 10^6) \times (1.6 \times 10^{-19})\ C/s$$

$$= 2.287 \times 10^{-13}\ A.$$

Example 2. *An ionization chamber is connected to an electrometer of capacitance* 0.5 *pF and voltage sensitivity of* 4 *divisions per volt. A beam of α-particles causes a deflection of* 0.8 *divisions. Calculate the number of ion pairs required and the energy of the α-particles. Given that* 1 *ion pair requires energy of* 35 *eV and* $e = 1.6 \times 10^{-19}$ *coulomb.*

Sol. Voltage sensitivity of electrometer = 4 divisions/volt.

$$\therefore \quad \left.\begin{matrix}\text{Voltage required to produce} \\ \text{a deflection of 0.8 divisions}\end{matrix}\right\} = \frac{0.8}{4}\ \text{volt} = 0.2\ \text{volt}$$

$$Q = CV = (0.5 \times 10^{-12}) \times 0.2 \ [\text{since } C = 0.5\ pF = 0.5 \times 10^{-12}\ F]$$

$$= 10^{-13}\ C.$$

$$\therefore \quad \text{No. of ion pairs required} = \frac{10^{-13}}{1.6\times10^{-19}} = 6.25\times10^5$$

1 ion pair requires 35 *eV*.

$$\therefore \quad \text{Total energy required} = 35 \times (6.25 \times 10^5)\ \text{eV}$$

$$= 21.88\ \text{MeV}.$$

29.4 Solid-state Detectors

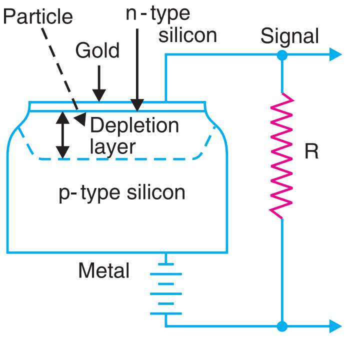
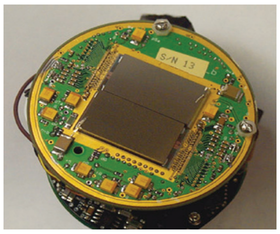


Fig. 29.5

A *p-n* junction, used as a particle detector, is shown in Fig. 29.5. It consists of a *p-n* junction between *p*-type and *n*-type silicon. Contact is made with the *n*-type silicon layer by a thin evaporated film of gold. In order to minimise the current flowing in the detector, when no radiation is striking it, a reverse biased diode is always used. The positive (reverse) bias applied to the gold film will push all the positive charge carriers away from the junction and produce a *depletion layer*, indicated in the figure. The depletion layer contains almost no carriers of either sign. When an energetic charged particle travels through the depletion layer, its interaction with the electrons in the crystal, produces electron-hole pairs. There is an electron-hole pair for every 3.5 *eV* (in *Si*) of energy lost by the charged particle. The electrons and holes are swept away by the applied electric field and registered as a voltage pulse over the resistor *R*. The number of charge carrier pairs produced in a semiconductor material is approximately 10 times as large as the number of ion pairs produced in a gas ion chamber *i.e.*, the energy extended per pair is about 3.5 *eV* in silicon, compared to about 30 *eV* for gases. The voltage pulse will therefore be about 10 times larger. Hence this detector has much better *energy resolution* than other radiation detectors. In solid state detectors for charged particles, silicon has been used most because of its low intrinsic conductivity. This means that the detector can be operated at room temperature without excessive leakage current. For gamma ray work, germanium detectors are much better than silicon because of the larger density of germanium.


Solid-State Detector.

29.5 Proportional Counter

The proportional counter consists of a cylindrical gas filled tube with a very thin central wire which serves as the anode (Fig. 29.6). The outer cylinder serves as a cathode. In the case of the simple ionization chamber, the pulse height generated by an event is proportional to the intensity of the beam. But because of the comparatively low applied voltages, the current produced is always very small.

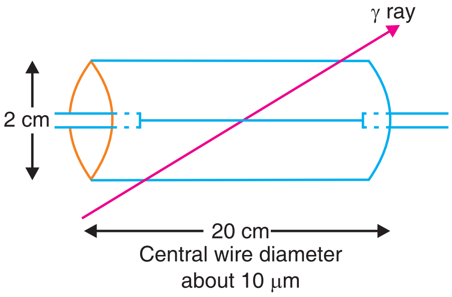


Fig. 29.6

If the voltage applied to an ionization chamber is increased past a certain value, the electrons acquire enough energy while moving toward the anode to create further ion pairs along the way. The resulting avalanche of secondary electrons that reaches the anode may represent a multiplication factor of as much as 1,000 with a correspondingly larger output pulse. In a certain range of applied voltages (See Fig. 29.7), the pulse size is proportional to the original number of ion pairs, and the device is called a proportional counter.

Since the central wire is very thin and the p.d. fairly large, the electric field $E = dV/dr$ at a distance r from the centre is very high. If b is the radius of the cylinder and a the radius of the wire, the radial field E at a distance r from the centre is given by

$$E = \frac{V}{r \log_e (b/a)}$$

where V is the positive voltage of the central wire relative to the outer cylinder. Thus, in a proportional counter, the field strength near the wire is very great. Hence electrons travelling towards the wire are rapidly accelerated when near it, and produce additional electrons in that region due to the phenomena of ionization by collision. This process is called the *gas multiplication.*

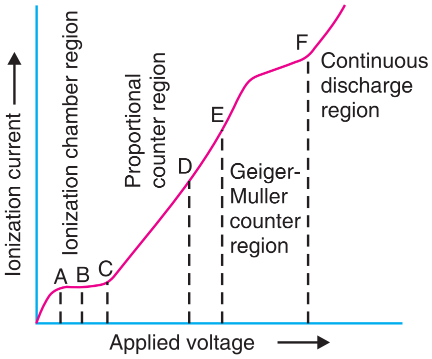


Fig. 29.7

The complete voltage pulse characteristics of this type of tube are shown in Fig. 29.7. The main regions used for measurements are : (1) The ionization chamber region *AB* (2) the proportional counter region *CD* (3) the Geiger-Muller region *EF.* After the point *F*, the tube becomes a simple discharge tube in which the current is produced even after the ionization event has ceased. Like the ionization chamber, the proportional counter gives single pulses of height proportional to the ionizing power of the radiation.

Example. *It is required to operate a proportional counter with a maximum radial field of* 10^7 Vm^{-1}. *What is the applied voltage required if the radii of the wire and tube are* 0.002 *cm and* 1 *cm respectively*?

Sol. Radial field = $E = \dfrac{V}{r \log_e (b/a)}$

Radial field at the wire surface $E = \dfrac{V}{a \log_e (b/a)}$

$$\therefore \qquad 10^7 = \frac{V}{(2\times10^{-5})\,2.302 \log_{10}\left(10^{-2}/2\times10^{-5}\right)}$$

$$\text{or} \qquad V = 1242 \text{ volts.}$$

29.6 Geiger-Muller Counter

It consists of a metal chamber *C* containing air or some other gas at a pressure of about 10 cm of Hg. A fine tungsten wire (*W*) is stretched along the axis of the tube and is insulated from it by ebonite plugs *EE* (Fig. 29.8). The wire is connected to the positive terminal of a high tension battery (about 1000 to 3000 volts) through a high resistance *R* (about 100 megohms) and the negative terminal is connected to the chamber *C*. The D.C. Voltage is kept slightly less than that which will cause a discharge between the electrodes.

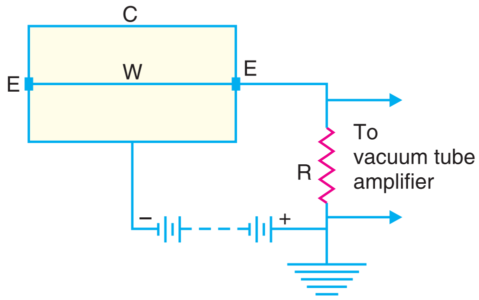
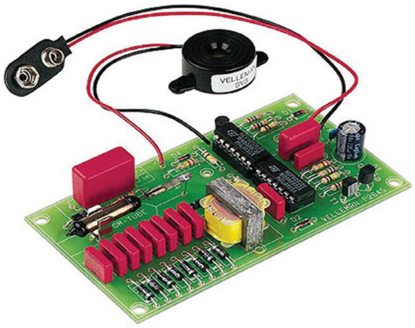


Fig. 29.8



Geiger-Muller Counter.

When an ionizing particle (say an α particle) enters the counter, ionisation takes place and a few ions are produced. If the applied P.D. is strong enough, these ions are multiplied by further collisions. An avalanche of electrons moves towards the central wire and this is equivalent to a small current impulse which flows through the resistance *R*. The critical potential is lowered momentarily, causing a sudden discharge through the resistance *R*. The p.d. thus developed across *R* is amplified by vacuum tube circuits and is made to operate a mechanical counter. In this way single particles can be registered. The sudden pulse of discharge sweeps away the ions from the chamber and the counter is ready to register the arrival of the next particle.

The voltage characteristics of a Geiger-Muller counter are shown in Fig. 29.9. This is a plot of the counting rate against the counter potential with a radioactive source placed near the counter. It is seen that there is a threshold below which the tube does not work. This can be several hundred volts. As the applied potential is increased, the counting begins and rises rapidly to a flat portion of the curve called the *plateau.* This is the region of

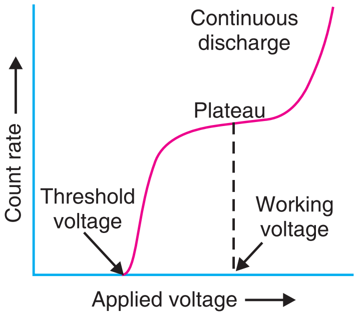


Fig. 29.9

the counter operation where the counting rate is, more or less, independent of small changes in p.d. across the tube. Beyond the plateau the applied electric field is so high that a continous discharge takes place in the tube as shown in Fig. 29.9 and the count rate increases very rapidly. It does not require any ionizing event for this to happen so that the tube must not be used in this region.

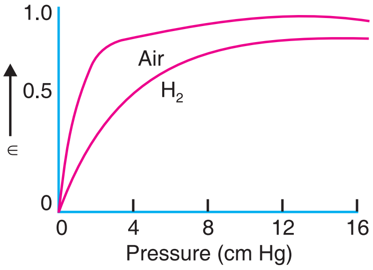


Fig. 29.10

The *efficiency of the counter* is defined as the ratio of the observed counts/sec. to the number of ionizing particles entering the counter per second. Counting efficiency is defined as the ability of its counting, if at least one ion pair is produced in it.

Counting efficiency = $\in = 1 - e^{slp}$ where s = specific ionization at one atmosphere; p = pressure in atmospheres and l = path length of the ionization particle in the counter. The efficiency ε of a *GM* counter, as a function of pressure for air and hydrogen, is illustrated in Fig. 29.10.

The counter set-up is portable (with the transistorised electronics) and serves best for mineral prospecting, apart from its several other applications in cosmic ray work. A virtue of the Geiger counter is that the pulse height is constant over a range of applied voltages, as in Fig. 29.9. So the power supply does not have to be precisely regulated as it does for a proportional counter. Also, the pulses are several volts in height, which makes amplifiers unnecessary.

Disadvantages of the Geiger counter are: (*i*) it is insensitive for a period of 200 to 400 µs following each pulse, which prevents its use at very high counting rates. (*ii*) it cannot provide information about the particle or photon causing a pulse.

EXAMPLE. *A self-quenched G-M counter operates at* 1000 *volts and has a wire diameter of* 0.2 *mm. The radius of the cathode is* 2 *cm and the tube has a guaranteed lifetime of* 10^9 *counts. What is the maximum radial field and how long will the counter last if it is used on an average for* 30 *hours per week at* 3000 *counts per minute? Consider 50 weeks to a year.*

SOL. The radial field at the central wire is

$$E_{\max} = \frac{V}{r\log_e(b/a)} = \frac{1000}{0.0001 \times 2.3026 \log_{10}\left(\dfrac{2\times10^{-2}}{10^{-4}}\right)}$$

$$= 1.89 \times 10^6 \text{ volts/metre.}$$

If the lifetime of the tube is N years, the total number of counts recorded will be

$$N \times 50 \times 30 \times 60 \times 3000 = 2.7 \times 10^8\, N$$

$$\therefore \quad 2.7 \times 10^8 \times N = 10^9$$

$$\text{or} \quad N = 3.7 \text{ years.}$$

29.7 The Wilson Cloud Chamber

Principle. If there is a sudden expansion of saturated vapour in a chamber, supercooling of the vapour occurs. Tiny droplets will be formed by condensation over the dust particles present in the chamber. If, therefore, we have completely dust-free and saturated air, and if it is suddenly allowed to expand and thereby cool, condensation will not take place. But if ions are available in the chamber during the expansion, they serve as nuclei for condensation. Hence, if an ionising particle passes through the chamber during an expansion, ions are produced along its path and droplets condense on these ions. Hence the "track" of the particle becomes visible.

Description. The apparatus consists of a large cylindrical chamber A, with walls and ceiling made of glass (Fig. 29.11). It contains dust-free air saturated with water vapour. P is a piston

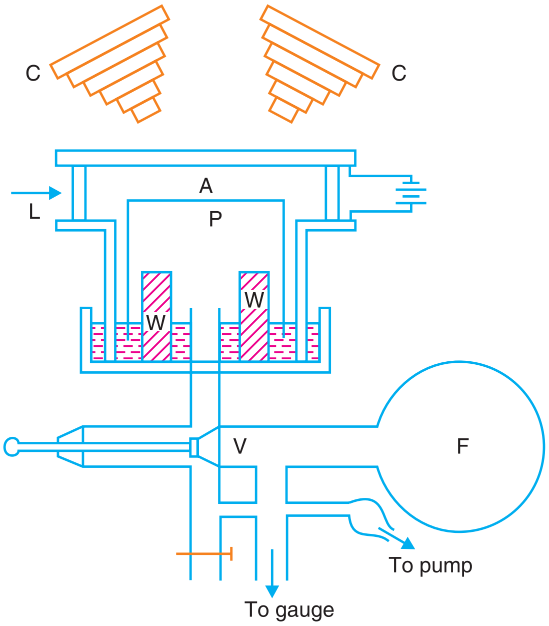


Fig. 29.11

working inside the chamber. When the piston moves down rapidly, adiabatic expansion of the air inside the chamber takes place. The piston is connected to a large evacuated vessel *F* through a valve *V*. When the valve is opened, the air under the piston rushes into the evacuated vessel *F*, thereby causing the piston to drop suddenly. The wooden blocks *WW* reduce the air space inside the piston. Water at the bottom of the apparatus ensures saturation in the chamber. The expansion ratio can be adjusted by altering the height of the piston.

As soon as the gas in the expansion chamber is subjected to sudden expansion, the ionizing particles are shot into the chamber through a side window. A large number of extremely fine droplets are formed on all the ions produced by the ionizing particles. These droplets form a track of the moving ionizing particles. At this stage, the expansion chamber is profusely illuminated by a powerful beam of light *L*. Two cameras *CC* are used to photograph the tracks. The process of expansion, shooting of the ionising particles into the expansion chamber, illuminating the chamber and clicking the camera must all be carried out in rapid succession in order to get satisfactory results.

Wilson Cloud Chamber

The ionising agent can be easily identified from its path in the cloud chamber. α-particles, being comparatively massive, go straight and their paths are *thick, straight* and *sharply defined.* β-particles being lighter, are easily deflected by collision and their paths are thin and *crooked.* The cloud chamber has led to the discovery of many elementary particles like positron, meson, etc.

Advantages. (1) Cloud chambers can be used to study the variation of specific ionisation along the track of a charged particle and the range of such particles.

(2) The sign of the electric charge and the momentum p of the particle can be determined if the chamber is placed in a strong magnetic field. Let a particle of mass m and charge q move with a velocity v perpendicular to the direction of the magnetic field of flux density B. The particle will be forced by the field to follow a circular path of radius R. The magnetic force Bqv is exactly balanced by the centrifugal force mv^2/R.

Thus $Bqv = Mv^2/R$ or $mv = p = BRq$.

The K.E. of the particle can be calculated, if the rest mass energy m_0c^2 of the particle is known, by the relation,

$$K.E. \;=\; E_k = \sqrt{\left[p^2c^2 + \left(m_0c^2\right)^2\right]} - m_0c^2$$

Limitations. (*i*) One is not always sure of the sense of track photographed.

(*ii*) The range of the particle may exceed the dimensions of the chamber so that the whole track is not photographed.

(*iii*) There remains a certain amount of uncertainty about the nature of the nuclei constituting the arms of the forked tracks.

29.8 Diffusion Cloud Chamber

The disadvantage of the cloud chamber lies in the fact that it needs a definite time to recover after an expansion. Hence it is not possible to have a continuous record of events taking place in the chamber. This difficulty was removed by the introduction of the diffusion cloud chamber.

The outline of the apparatus is shown in Fig. 29.12.

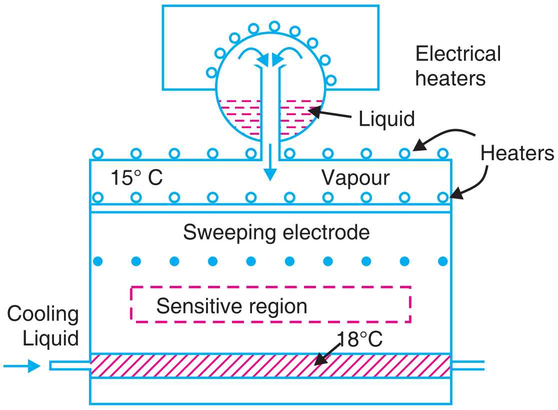
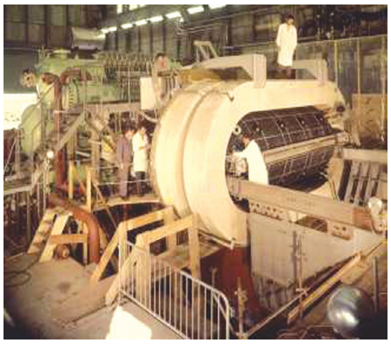


Fig. 29.12

It consists of a chamber containing a heavy gas which is kept warm at the top and cold at the bottom. Thermal gradient is maintained between the bottom and top of the chamber by external heating or cooling. The liquid (methyl alcohol) vaporises in the warm region, where the vapour pressure is high. The vapour diffuses downwards continuously where the vapour pressure is low and condensation takes place. In a region near the base, the supersaturation factor is high and condensation takes place around the available ions. The chamber remains continuously sensitive to ionizing particles until the supply of volatile liquid is exhausted. The system is illuminated by a strong source of light and the track of the particle is photographed by camera.

29.9 Bubble Chamber

Principle. We know that normally the liquid boils with the evolution of bubbles of vapour at the boiling point. If the liquid is heated under a high pressure to a temperature well above its normal boiling point, a sudden release of pressure will leave the liquid in a *superheated state.* If an ionising particle passes through the liquid within a few milliseconds after the pressure is released, the ions left in the track of a particle act as condensation centres for the formation of vapour bubbles. The vapour bubbles grow at a rapid rate and attain a visible size in a time of the order of 10 to 100μs. Thus in a bubble chamber, a vapour bubble forms in a superheated liquid, whereas in a cloud chamber, a liquid drop forms in a supersaturated vapour. Thus an ionising particle passing through the superheated liquid leaves in its wake a trail of bubbles which can be photographed.



Bubble Chamber.

A schematic diagram of a liquid hydrogen bubble chamber, operating at a temperature of 27 *K* is shown in Fig. 29.13. A box of thick glass walls is filled with liquid hydrogen and connected to the expansion pressure system. To maintain the chamber at constant temperature, it is surrounded by liquid nitrogen and liquid hydrogen shields. High energy particles are allowed to enter the chamber from the side window *W*. A sudden release of pressure from the expansion valve is followed by light flash and camera takes the stereoscopic view of the chamber.

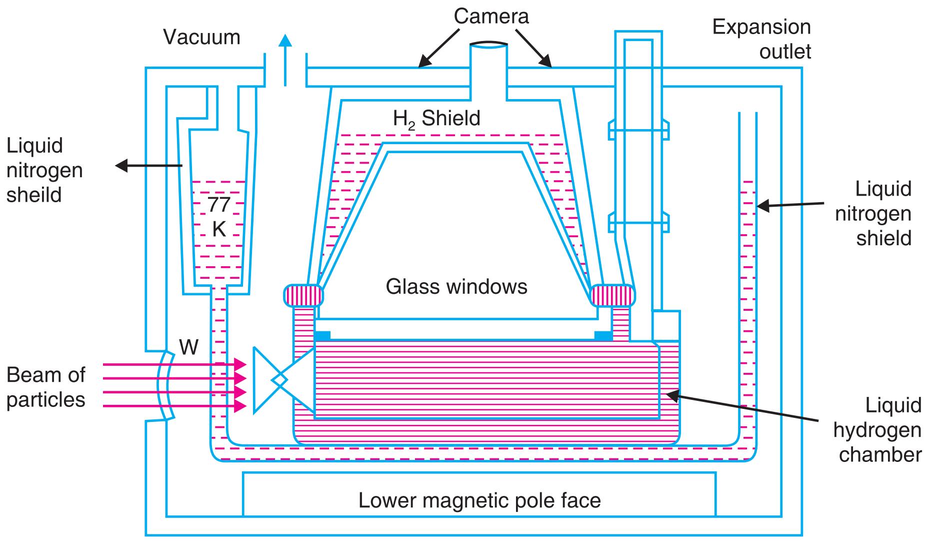


Fig. 29.13

The incoming beam triggers the chamber. The charge of the tracks can be identified by the direction of their curvature in the magnetic field applied over the bubble chamber. From the curvature and length of the track, the momentum and energy of the particle can be found. The bubble chamber is used to *study particle interaction* and to *detect very high energy particles.*

Advantages. (1) The density of a liquid is very large when compared to that of a gas of even high pressure. Hence the chances of collision of a high energy particle with a molecule of the liquid are very much greater. Consequently there is a greater chance of their track being recorded. So the chances of recording events like cosmic ray phenomena are improved when compared with cloud chambers.

(2) The bubbles grow rapidly and as a result the tracks are not likely to get distorted due to convection currents in the liquid.

(3) The bubble chamber is sensitive even to particles of low ionising power.

29.10 Spark Chamber

A spark chamber consists of a set of conducting plates alternately connected to a source of high DC voltage (Fig. 29.14)

The chamber is filled with an inert gas. Sudden application of very high voltages to alternate plates, while the others are left at ground potential, results in very high electrical fields across the gaps. Electrical breakdown then occurs along the trails of ions. So the trajectory of a given particle through the system is marked by a series of sparks. The spark trails are photographed stereoscopically. If the chamber is located in a magnetic field, the charge and momentum of the particle can be determined from the curvature of the track. Vidicons are frequently used in place of photography.

The spark chamber has an important advantage over the bubble chamber in that it can be rendered sensitive for only a very short time. Suppose, for example, that in an intense beam from a high-energy particle accelerator a rather rare particle is produced. Scintillation or Cerenkov counters can signal the production of this kind of particle. The pulse voltage is then applied to the spark chamber. The track of this particular particle can thus be determined. There is little likelihood that one of the much more numerous *background* particles will cause another spark track in the short time that the chamber is sensitive. Thus rare events and processes can be studied.

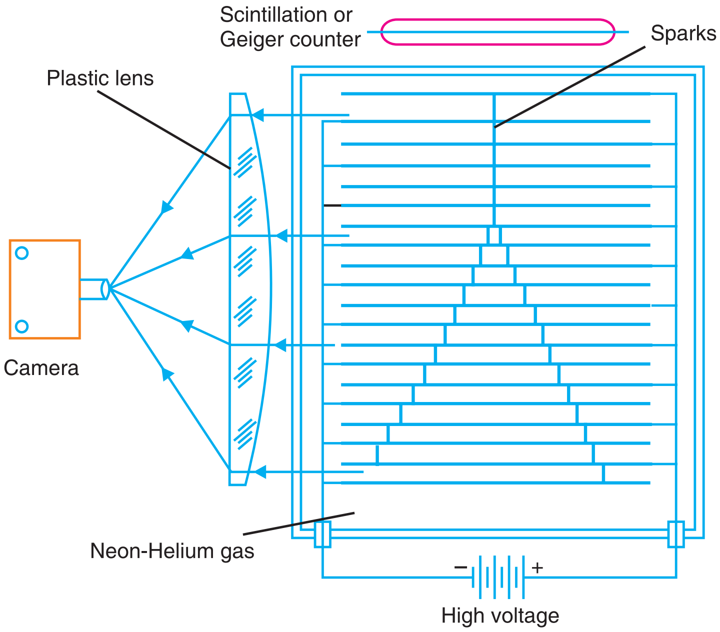


Fig. 29.14

29.11 Nuclear Emulsions

Nuclear Emulsion is the Detector with the best Spatial Resolution.

It has been known for a long time that charged particles affect a photographic plate. A heavy charged particle traversing a photographic emulsion produces a latent image of its track. The track is revealed when the plate is developed. Ordinary optical photographic emulsions are not suitable for quantitative studies with nuclear radiations. The sensitivity of such emulsions is low. Further, the tracks due to charged particles have non-clear range because the developed crystal grains are large and widely spaced. The composition of the emulsion was changed so as to make it more suitable for the study of various ionising particles, such as α-particles, protons, mesons and even electrons. The nuclear emulsions differ from the optical emulsions in that they have considereably higher silver halide content and smaller grain size. In nuclear emulsion, the thickness is greater than that of optical emulsion (Table 29.1). The smaller the grain size, the more sensitive is the emulsion to ionising particles. Thus different commercially available emulsions, differing chiefly in grain size, can be used to discriminate between different particles.

TABLE 29.1 **Nuclear and Optical Emulsions**

Property	*Optical Emulsion*	*Nuclear Emulsion*
AgBr : Gelatin (mass)	47 : 53	80 : 20
AgBr : Gelatin (volume)	15 : 85	45 : 55
Grain size (micron)	1–3.5	0.1–0.6
Thickness („)	2–3	25–2000
Sensitivity to light	Very high	Poor
Response to α particle	Dense blackening	Individual tracks
” ” β ”	Moderate	Faint fog
” ” γ ”	Faint	Almost none

Advantages. (1) The emulsion is relatively light and cheap. Because of their lightness, they can be sent in balloons, spaceships etc., for high altitude cosmic ray experiments. The cosmic ray events once recorded can be studied by developing the exposed plates conveniently in the laboratory. Emulsions were widely employed in cosmic ray studies and led to the discovery of the π and K mesons.

(2) The high density of the emulsion gives it a stopping power about a thousand times that of standard air. Unstable high energy particles are brought to rest in the emulsions and their decay schemes can thus be studied.

(3) The emulsion is continuously sensitive and is consequently always available to record an event. But the cloud chamber is sensitive only for a fraction of a second after an expansion and remains ineffective for several seconds between successive expansions.

Limitations. (1) The main drawback of nuclear emulsions is that their sensitivity and thickness are affected by temperature, humidity, age of the emulsions before development and the conditions under which they are developed. The scanning of the plates and the analysis of the tracks obtained are also laborious, when done manually.

(2) As compared with cloud chamber, photographic emulsions have the following drawback. It is very difficult to determine the sign of the charge and the momentum of a particle from observations of the curvature of the path in a strong magnetic field. The actual path-length in the emulsion is small compared to that in a cloud chamber.

29.12 The Scintillation Counters

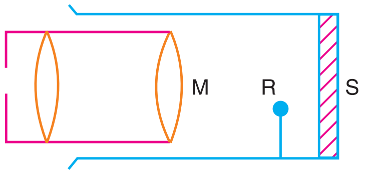


Fig. 29.15

Scintillation Counter.

One of the earliest methods of radiation detection was the spinthariscope (Fig. 29.15). It consists of a small wire, the tip of which is dipped in Radium bromide (R) or any other radioactive salt. It is placed in front of a zinc sulphide screen S and viewed through a microscope. When an α or β-particle falls on the zinc sulphide screen, they produce light flashes which can be seen by a microscope (M) in a dark room. The visible luminescence excited in zinc sulphide by α-particles was used by Rutherford for counting the particles. The process of counting these scintillations through a low power microscope is a tedious one and the limitations of observation with the eye restrict the counting rate to about 100 per minute. This process, whereby the energy of the particle is converted to light, is the basis of scintillation counter.

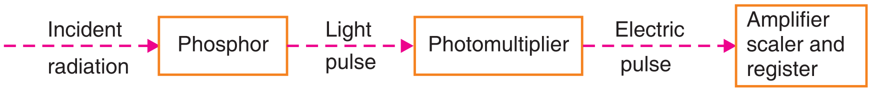


Fig. 29.16

The main parts of a scintillation counter are shown in Fig. 29.16. The atoms of the phosphor are excited or ionised by the energy loss of an impinging α, β or γ ray. When the atoms return to their ground states, photons are emitted, in the blue and ultraviolet regions of the optical spectrum. The phosphor is optically coupled to the envelope of a photomultiplier tube. The photons strike the photocathode, causing the ejection of photo-electrons (Fig. 29.17). As these photo-electrons leave the photocathode, they are directed by a focusing electrode to the first multiplier electrode or *dynode.* This electrode has the property of emitting three, four or five electrons for every single electron which strikes its surface. There may be from 10 to 14 such multiplier stages in a given tube. Hence, from the emission of one single electron from the cathode, a burst of one million electrons may impinge on the final stage in the tube (the anode). The output pulse from the photomultiplier is fed to a pulse amplifier followed by a scaler circuit.

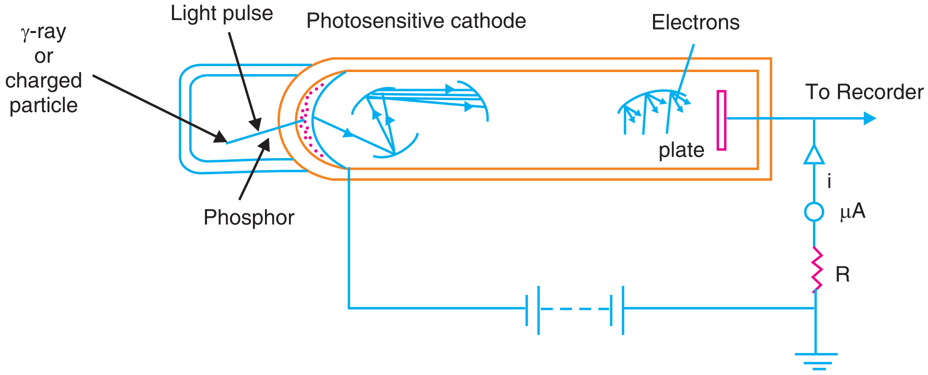


Fig. 29.17

29.13 Cerenkov Counter

Cerenkov Counter.

When a charged particle moves through a transparent dielectric medium with a velocity greater than the velocity of light in that medium, a cone of light waves is emitted. These light waves are known as *Cerenkov radiation.* In a dielectric medium of refractive index *n*, photons move with a velocity *c/n.* Consider a charged particle moving with a velocity *v* through the dielectric medium. Fig. 29.18 shows the interactions when $v > c/n$. The envelope of the radiation is a cone of half angle θ with the particle at its apex, where

$$\sin\theta = \frac{(c/n)t}{vt} = \frac{c}{nv}.$$

The angle θ of the cone of radiation depends on the speed *v* of the particle. When a beam of fast charged particles moves through a medium such as glass or transparent plastic, the radiation can be detected and the angle θ measured. Thus the speed of the particle can be determined.

Fig. 29.19 shows a Cerenkov detector designed by Marshall. A collimated beam of pions is allowed to pass along the axis of a large hemispherical perspex lens. The Cerenkov radiation radiated in this lens is reflected by the cylindrical mirror M_1 on to one of the plane mirrors M_2, M_3 and then on to the cathodes of the photomultiplier tubes (*IP* 28). The two photomultipliers operate in coincidence. The position of the radiator lens is adjusted so that the Cerenkov light cone is correctly focused on to the phototubes to give maximum coincidence counting rate.

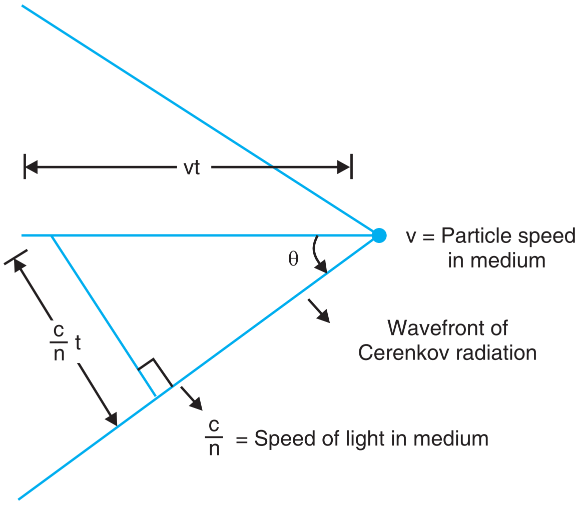

vt
v = Particle speed in medium
θ
$\frac{c}{n}$ t
Wavefront of Cerenkov radiation
$\frac{c}{n}$ = Speed of light in medium

Fig. 29.18

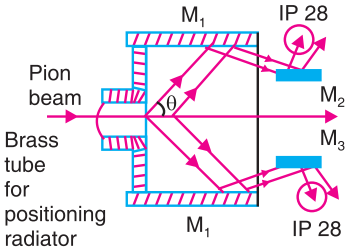


Fig. 29.19

The Cerenkov detector is used for the detection of high energy particles. It has a high efficiency and a high counting rate.

EXERCISE

1. Describe the construction and working of an ionisation chamber.
2. Describe a G.M. counter and explain its working as a particle detector.
3. Explain fully the working of a cloud chamber. How is it used to determine the energy of a particle passing through it?
4. What are the constructional features of the bubble chamber? How does the instrument work? Enumerate its advantages over the cloud chamber.
5. Give a brief account of the nuclear emulsion technique in the study of ionising radiations.
6. Give an account of the mode of operation of a scintillation counter and describe how it may be utilised in the study of nuclear reactions.
7. Write a brief essay on the various techniques which are used for the detection of elementary particles indicating the advantages and the limitations pertaining to each method.
8. A G.M. counter wire collects 10^8 electrons per discharge. When the counting rate is 500 counts/min, what will be the average current in the circuit? [**Ans.** 1.33×10^{-10} *A*]
9. A Geiger-Muller counter collects 10^7 electrons per discharge. The average current in the circuit is 1.333×10^{-11} amp. Find the counting rate per minute. $e = 1.6 \times 10^{-19}$ coulomb. [**Ans.** 500]
10. A high energy singly ionised particle leaves a track in a cloud chamber placed in a magnetic field of flux density 2.5 Wb/m^2 and the radius of curvature of the track is 200 cm. Calculate the energy of the particle.

    [**Hint.** Energy $= E = pc = BRqc = 2.5 \times 2 \times 1.602 \times 10^{-19} \times (3 \times 10^8)\, J = 240\, MeV$]
11. Calculate the ionisation current produced by 3 *MeV* deuterons passing through a gas at 1000 per sec. assuming that 25 *eV* is required to produce an ion pair. [**Ans.** 1.923×10^{-11} *A*]

12. An ionisation chamber exposed to a beam of α-particles registers a current of 3.85×10^{-13} *A*. On the average, 20 α-particles enter the chamber per second. Assuming that the production of an ion pair in the chamber involves the expenditure of 35 *eV* of energy, calculate the energy of the α-particles. [**Ans.** 4.21 *MeV*]

13. An organic-quenched G.M. tube operates at 1000 volts and has a wire diameter of 0.2 mm. The radius of the cathode is 2 cm, and the tube has a guaranteed life-time of 10^9 counts. What is the maximum radial field and how long will the counter last if it is used on the average for 60 hours per week at 2000 counts per minute? [**Ans.** $E_{max} = 1.89 \times 10^6\ V\ m^{-1}$; $N = 2.77$ years]

14. Measurements on the track of a particle carrying a charge *e* in a detector give the value of 330 *MeV/c* for its momentum and 215 *MeV* for its kinetic energy. Find the rest mass of the particle. If the detector is kept in a magnetic field of 1 *T* and the particle moves normal to the field, what is the radius of curvature of its track in the detector? [**Ans.** 145.8 *MeV*; 1.1 *m*]

15. What is the meaning and significance of quenching in a *GM* counter?

Ans. The process of removing all the ions from the chamber due to continuous discharge and making it ready for fresh event is called *quenching*. This is achieved by using quenching agents like alcohol or halogens.

Quenching Process. When a G.M. tube operates in Geiger region, the secondary electrons increase the current pulse by further ionisation of gas molecules. The object of counter is to produce a single pulse due to entry of a single particle. The tube should not then give any succeeding spurious pulses but should recover as quickly as possible to be in the state to record the next entering particle. Therefore, it is desirable to prohibit the production of spurious pulses following the main required pulse due to a single particle entry. The process of prohibiting the undesirable secondary pulses is called the *quenching*.

Self-quenching of GM-counter. A typical GM-counter contains 90% Ar and 10% ethyl alcohol by weight. The ionisation potential of alcohol (11.3 eV) is lower than that of argon (15.7 volt). As a result, the inert gas ions get neutralized by transferring their charge to alcohol ions during their long path towards cathode. The alcohol ions produced, capture electrons from the cathode and are neutralized. Hence there is no multiple pulsing and the discharge is quenched soon (fraction of a milli-second) after the initial ionisation.

16. What is meant by the *resolving time* of a GM-counter?

Ans. Fig. 29.20 gives the three time-intervals—*dead time, recovery time* and *resolving time*—of operation of a *GM*-counter.

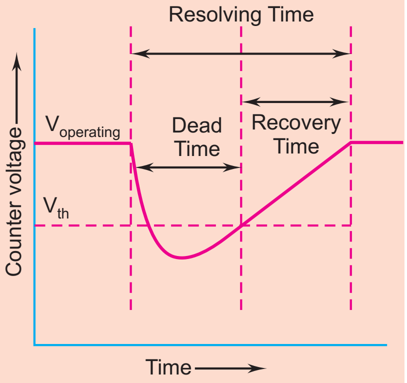


Fig. 29.20

(*i*) **Dead time.** The time, during which the counter is incapable of responding to a second ionizing event, is known as *dead time* (~ 200 μ sec). It is called *dead time* because during this period the counter is insensitive (dead) to further pulses. After the initial ionisation, the cloud of slow moving massive positive ions formed over the anode wire effectively lowers its potential relative to cathode. Thus the G.M counter is dead for about 200 μs after each ionising event.

(*ii*) **Recovery time.** As the positive ion-sheath moves further toward the cathode, the counter sensitivity gradually returns to its original value and the second pulse can now be recorded. This time, during which pulses of reduced size are poduced, is known as *recovery time*.

(*iii*) **Resolving time.** The dead time plus recovery time corresponds to the time which positive ion sheath takes to reach the cathode, and is known as the *resolving time* of the counter. This is the time between just recorded pulses.

Expression for resolving time. Assume that a counter with a resolving time τ responds at a rate *n* counts per unit time when exposed to *N* particles per unit time. In unit time, the total insensitive time is $n\tau$. The number of counts missed is $Nn\tau$.

The number of counts missed is the error in counting.

$$Nn\tau = N - n \qquad ...(1)$$

$$\therefore \qquad N = n/(1 - n\tau) \qquad ...(2)$$

From Eq. (2), we can find the *true* counting rate N from the *observed* counting rate n, if τ is known.

Determination of τ. By measuring separately the counting rates n_1' and n_2' with two radioactive sources 1 and 2 of nearly equal strength, τ can be determined. Let n_t' denote the total counting rate when both sources are placed in the same position as before and n_b be the background rate (*i.e.,* with no source).

$$\therefore \qquad n_1 = n_1' - n_b; \quad n_2 = n_2' - n_b; \quad n_t = n_t' - n_b \qquad ...(3)$$

Here, n_1, n_2, n_t are the above counting rates corrected for background.

If N_1, N_2 and N_t be the true rates of arrival of particles in the counter in the above three situations, we have, using Eq. (2),

$$N_1 = \frac{n_1}{1 - n_1\tau}; \quad N_2 = \frac{n_2}{1 - n_2\tau}; \quad N_t = \frac{n_t}{1 - n_t\tau} \qquad ...(4)$$

Also, $$N_t = N_1 + N_2 \qquad ...(5)$$

Assuming the values of terms involving τ^2 to be negligible, we have, solving for τ

$$\tau = \frac{n_1 + n_2 - n_t}{2n_1n_2} \qquad ...(6)$$

17. In a scintillation counter, the scintillator generally used is

(*a*) Anthracene (*b*) Argon (*c*) Camphor (*d*) Neon (*Barkatullah University, 2011*)

[**Ans.** (*a*)]

18. The wire in a GM counter collects 10^{10} electrons per discharge. If the count rate is 1000/minute, calculate the average current in the circuit. (charge on electron is 1.6×10^{-19} C). (*Andhra University, 2010*).

Ans. Counting rate = 1000 counts/min.

The wire collects 10^{10} electrons per discharge.

$\therefore$ Total number of electrons collected in one min. $= n = 1000 \times 10^{10} = 10^{13}$

$$\text{Charge/min} = n\,e = 10^{13} \times (1.6 \times 10^{-19}) = 1.6 \times 10^{-6} \text{ coul/min.}$$

$$\text{Charge/second} = \frac{1.6\times10^{-6}}{60} = 2.66 \times 10^{-8} \text{ Amp.}$$

$$\therefore \qquad \text{Average current} = \text{charge/second} = 2.66 \times 10^{-8} \text{ Amp.}$$

CHAPTER
30

PARTICLE ACCELERATORS

AT A GLANCE

30.1 Introduction
30.2 Van de Graaff Generator
30.3 The Linear Accelerator
30.4 The Cyclotron
30.5 The Synchrocyclotron
30.6 The Betatron
30.7 The Synchrotrons
30.8 The Proton Synchrotron (Bevatron, Cosmotron)

30.1 Introduction

A particle accelerator is a device for increasing the K.E. of electrically charged particles. Methods of acceleration can be classed into three groups : *direct field*, *inductive*, and *resonance.* According to the shape of the path of the particles, accelerators are classified as *linear* and *cyclic.* In linear accelerators, the paths of particles are approximately straight lines; in cyclic accelerators, they are circles or spirals. (*i*) In a direct field linear accelerator, a particle passes only once through an electric field with a high p.d. set up by electrostatic generators. (*ii*) The only accelerator of the inductive type is the betatron. (*iii*) In *magnetic resonance* accelerators, the particle being accelerated repeatedly passes through an alternating electric field along a closed path, its energy being increased each time. A strong magnetic field is used to control motion of particles and to return them periodically to the region of the accelerating electric field. The particles pass definite points of the alternating electric field approximately when the field is in the same phase ("in resonance"). The simplest resonance accelerator is the cyclotron.

The Cockcroft-Walton tension multiplier

The principle of this instrument is that of cascade voltage rectification or voltage multiplication by suitable arrangement of high tension rectifier tubes and charging capacitors, as shown in

Fig. 30.1. It consists of a high voltage transformer T and two banks of identical type series capacitors alternately connected by identical type high voltage rectifiers. The high voltage step-up transformer T supplies the A.C. voltage to be rectified by H.T. valves V_1, V_2, V_3, V_4 etc. If in the first half-cycle the valve V_1 rectifies the A.C., it charges the capacitor C_1 to the maximum A.C. voltage V_o. In the next half-cycle, V_1 will not conduct, but V_2 will rectify and charge C_2 to a voltage $2V_0$. In this manner, the capacitors C_3, C_4 etc., will go on getting charged in series to higher and higher voltages. If a large number of such steps are used, a very high voltage D.C. supply will be available. The output of this supply will be approximately equal to nV_o where n is the number of H.T. valves used and V is the maximum A.C. voltage amplitude. This output voltage can be used to accelerate ions.

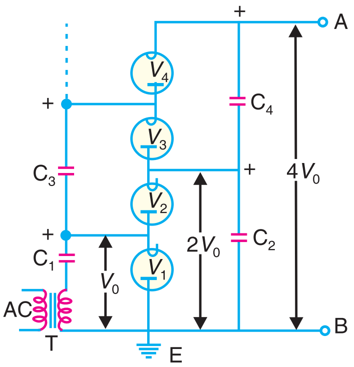
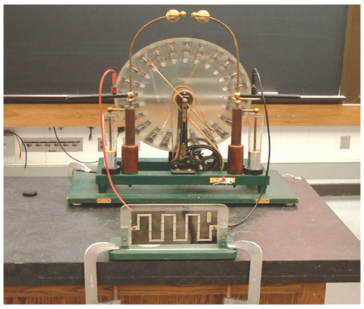


Fig. 30.1

30.2 Van de Graaff Generator

In this electrostatic accelerator, electric charge is carried continuously to an insulated hollow spherical conductor (S) and delivered to it. The potential of the hollow sphere can thus be gradually raised. When the potential has attained a sufficiently high value, it can be used to accelerate positive ions in the accelerator tube.



Van De Graaff Generator.

A well insulated endless belt B is rapidly rotated between two pulleys P_1 and P_2 by means of a motor [Fig. 30.2 (a)]. The upper pulley P_2 is placed inside a hollow spherical conductor S. Near the lower pulley P_1, there is a metallic comb C connected to a source of steady voltage of 50 kV. Positive charge is sprayed over the moving belt by the metallic comb C. The belt moving upward carries this charge into the hollow sphere. This positive charge, on entering the sphere S, induces an equivalent negative charge on its inside and a corresponding positive charge on its outside. The negative charge is sprayed by the comb C' on to the belt, thus neutralising the positive charge on it. The net result is that the positive charges are in effect transferred from the moving belt to the outside of the sphere. The belt, returning to the bottom, is charged again and delivers up the charge, when inside the sphere. Charges are thus accumulated and distributed on the surface of the sphere. As the charge on the sphere accumulates, its potential rises to a very high value.

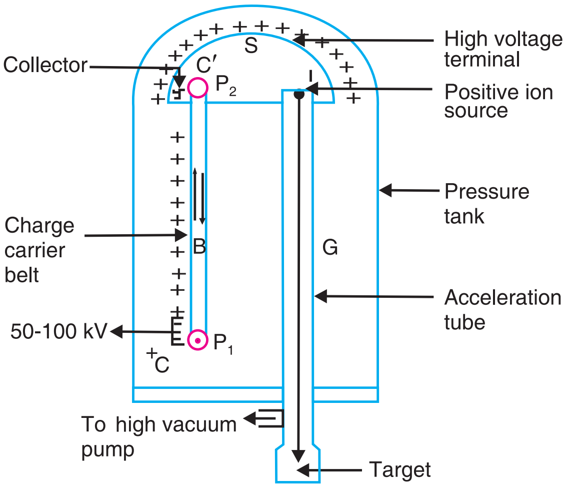


Fig. 30.2(a)

The charged particles which are to be accelerated are produced by a gas discharge ion source I placed inside the hollow sphere S over the accelerating tube G. The accelerating tube is evacuated and contains a number of insulated metal cylinders (C) arranged with short gaps between them [Fig. 30.2]. The H.T. is applied to the topmost cylinder near the ion source and the bottom-most cylinder is earthed. A potential gradient exists from the top to the bottom. The ions are accelerated at the gaps between the cylinders, where intense electric field exists. The highly accelerated ions are deviated by a suitable and adjustable magnetic field M and made to strike the target (T) at the bottom of the tube. By reversing the potential of the spray voltage, the Van de Graaff generator can also be used to accelerate electrons.

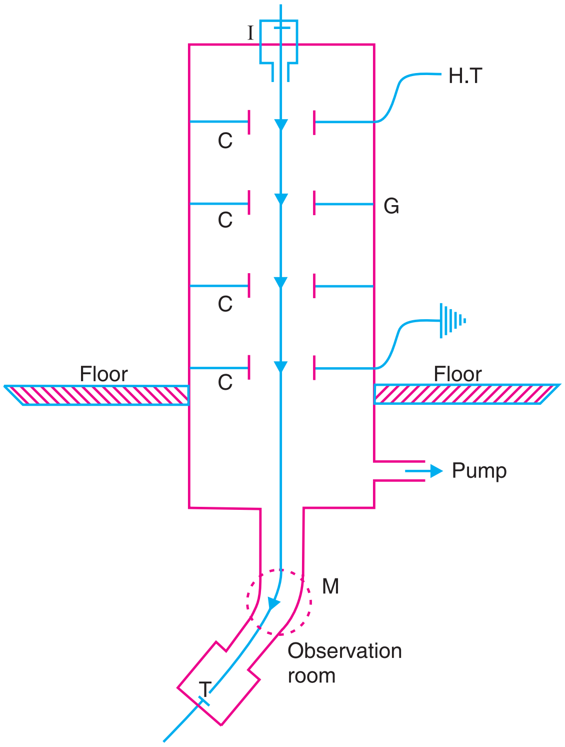
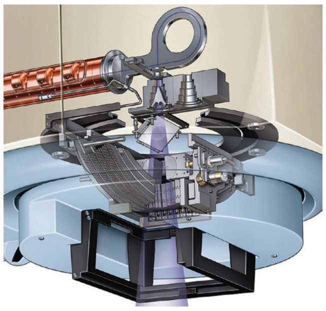


Fig. 30.2 (b)

30.3 The Linear Accelerator

Direct acceleration of particles by potentials above 10 million volts is a difficult problem due to insulation difficulties. For such high energies, acceleration of the particles is achieved in small successive steps. In such machines, the P.D. between different parts of the machine and between the machine and earth, is maintained low, compared with the P.D. corresponding to the ultimate energy acquired by the particles. One machine employing this method is the linear accelerator. In this machine, high energy particles are produced without employing high P.D.'s, by using the principle of *synchronous acceleration.*



Linear accelerator.

Fig. 30.3 shows the schematic diagram of a linear accelerator. It consists of a series of coaxial hollow metal cylinders or drift tubes 1, 2, 3, 4 etc. They are arranged linearly in a glass vacuum chamber. The alternate cylinders are connected together, the odd numbered cylinders being joined to one terminal and the even numbered ones to the second terminal of a H.F. oscillator. Thus in one-half cycle, if tubes 1 and 3 are positive, 2 and 4 will be negative. After half a cycle the polarities are reversed *i.e.*, 1 and 3 will be negative and 2 and 4 positive. The ions are accelerated only in the gap between the tubes where they are acted upon by the electric field present in the gaps. The ions travel with constant velocity in the *field-free space* inside the drift tubes.

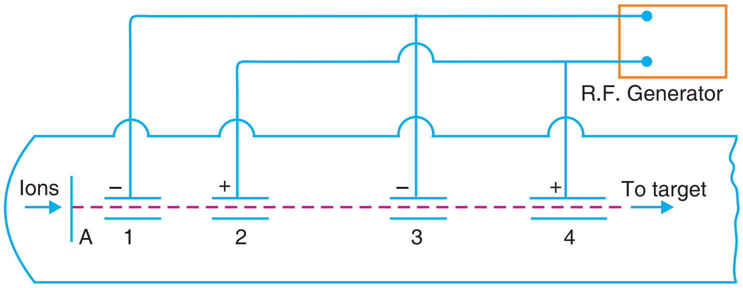


Fig. 30.3

Positive ions enter along the axis of the accelerator from an ion source through an aperture A. Suppose a positive ion leaves A and is accelerated during the half cycle, when the drift tube 1 is negative with respect to A. Let e be the charge and m the mass of the ion and V potential of drift tube 1 with respect to A. Then velocity v_1 of the ion on reaching the drift tube is given by

$$\frac{1}{2}mv_1^2 = Ve \text{ or } v_1 = \sqrt{\frac{2Ve}{m}}.$$

The length of the tube 1 is so adjusted that as the positive ions come out of it, the tube has a positive potential and the next tube (tube No. 2) has a negative potential, *i.e.*, the potentials change sign. The positive ion is again accelerated in the space between the tubes 1 and 2. On reaching the tube 2, the velocity v_2 of the positive ion is given by

$$\frac{1}{2}mv_2^2 = 2Ve \text{ or } v_2 = \sqrt{2}\sqrt{\frac{2Ve}{m}} = \sqrt{2}\,v_1.$$

This shows that v_2 is $\sqrt{2}$ times v_1. In order that this ion, on coming out of tube 2, may find tube 3 just negative and the tube 2 positive, it must take the same time to travel through the tube 2. Since $v_2 = \sqrt{2}\,v_1$, the length of tube 2 must be $\sqrt{2}$ times the length of tube 1. For successive accelerations in successive gaps the tubes 1, 2, 3, etc., must have lengths proportional to 1, $\sqrt{2}$, $\sqrt{3}$ etc. *i.e.*, $l_1 : l_2 : l_3$: etc. = 1 : $\sqrt{2}$, $\sqrt{3}$: etc.

Energy of the ion. If n = the number of gaps that the ion travels in the accelerator and v_n = the final velocity acquired by the ion, then

$$\left.\begin{array}{l}\text{Velocity of the ion, as it}\\ \text{emerges out of the } n^{th} \text{ tube}\end{array}\right\} = \sqrt{n}\sqrt{\frac{2Ve}{m}}.$$

$$\therefore \quad \text{K.E. acquired by the ion} = \frac{1}{2}mv_n^2 = nVe.$$

Thus the final energy of the ions depends upon (*i*) the total number of gaps and (*ii*) the energy gained in each gap.

The limitations of this accelerator are : (*i*) The length of the accelerator becomes inconveniently large and it is difficult to maintain vacuum in a large chamber. (*ii*) The ion current available is in the form of short interval impulses because the ions are injected at an appropriate moment.

EXAMPLE. *In a linear accelerator, proton accelerated thrice by a potential of* 40 *kV leaves a tube and enters an accelerating space of length* 30 *cm before entering the next tube. Calculate the frequency of the r. f. voltage and the length of the tube entered by the proton.*

SOL. Let v_1 and v_2 be the velocities of the proton on entering and leaving the accelerating space. Let e and m be the mass and charge of the proton respectively. Then

$$\frac{1}{2}mv_1^2 = 3 \times e \times 40000$$

$$\begin{aligned}\therefore \quad v_1 &= [(2 \times 3 \times 40000 \times (e/m)]^{1/2}\\ &= [2 \times 3 \times 40000 \times (9.578 \times 10^7)]^{1/2}\\ &= 4.794 \times 10^6\ ms^{-1}.\end{aligned}$$

$$\begin{aligned}\text{Similarly,} \quad v_2 &= \left[2\times4\times40000\times(9.578\times10^7)\right]^{1/2}\\ &= 5.536 \times 10^6\ ms^{-1}.\end{aligned}$$

Mean velocity while travelling the 0.3 m distance

$$= 5.165 \times 10^6\ ms^{-1}.$$

The time taken to travel 0.3 m is half the period ($T/2$) of the $r.f.$ voltage.

$$\therefore \qquad \frac{T}{2} = \frac{0.3}{5.165 \times 10^6} \text{ or } T = \frac{2 \times 0.3}{5.165 \times 10^6}$$

$$\therefore \qquad \left.\begin{array}{l}\text{frequency of the} \\ r.f.\text{ voltage}\end{array}\right\} = \frac{5.165 \times 10^6}{2 \times 0.3}$$

$$= 8.608 \times 10^6\ Hz = 8.608\ MHz.$$

The protons travel through the next tube for half a period with a velocity of $5.536 \times 10^6\ ms^{-1}$.

$\therefore$ length of the tube entered by the protons

$$L = 5.536 \times 10^6 \times \frac{1}{2 \times 8.608 \times 10^6} = 0.3216\ m$$

30.4 The Cyclotron

Construction. The cyclotron (Fig. 30.4) consists of two hollow semicircular metal boxes, D_1, D_2 called ''dees''. A source of ions is located near the mid-point of the gap between the ''dees''. The "dees" are insulated from each other and are enclosed in another vacuum chamber. The "dees" are connected to a powerful radio-frequency oscillator. The whole apparatus is placed between the pole-pieces of a strong electromagnet. The magnetic field is perpendicular to the plane of the "dees".

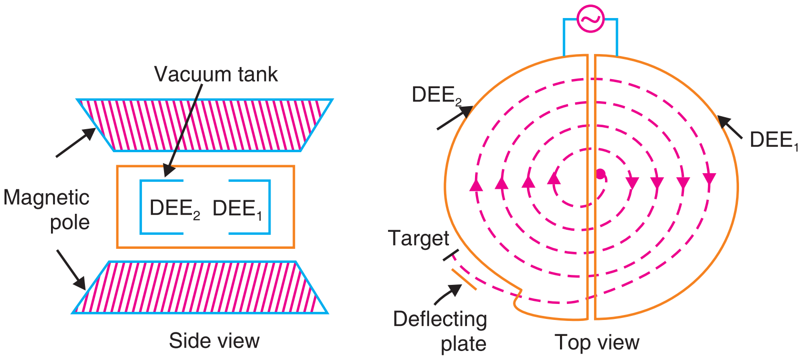
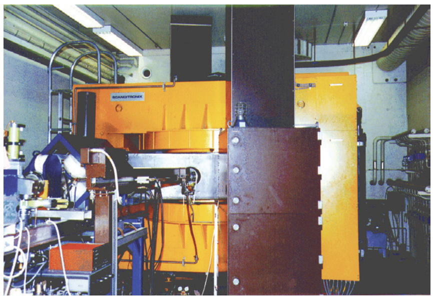


Fig. 30.4



Cyclotron.

Theory. Suppose a positive ion leaves the ion source at the centre of the chamber at the instant when the "dees" D_1 and D_2 are at the maximum negative and positive A.C. potentials respectively. The positive ion will be accelerated towards the negative dee D_1 before entering it. The ions enter the space inside the dee with a velocity v given by $Ve = \frac{1}{2}mv^2$, where V is the applied voltage and e and m are the charge and mass of the ion respectively. When the ion is inside the "dee" it is not accelerated since this space is field free. Inside the dee, under the action of the applied magnetic field, the ions travel in a circular path of radius r given by

$$Bev = mv^2/r \qquad \text{...(1)}$$

where B = the flux density of the magnetic field.

or $$r = mv/Be \qquad \text{...(2)}$$

The angular velocity of the ion in its circular path = $\omega = \dfrac{v}{r} = \dfrac{Be}{m}$ ...(3)

The time taken by the ion to travel the semicircular path = $t = \dfrac{\pi}{\omega} = \dfrac{\pi m}{Be}$ ...(4)

Suppose the strength of the field (B) or the frequency of the oscillator (f) are so adjusted that by the time the ion has described a semicircular path and just enters the space between D_1 and D_2, D_2 has become negative with respect to D_1. The ion is then accelerated towards D_2 and enters the space inside it with a greater velocity. Since the ion is now moving with greater velocity, it will describe a semicircle of greater radius in the second "dee". But from the equation $t = \pi m/Be$ it is clear that the time taken by the ion to describe a semicircle is independent of both the radius of the path (r) and the velocity of the ion (v). Hence the ion describes all semicircles, whatever be their radii, in exactly the same time. This process continues until the ion reaches the periphery of the dees. The ion thus spirals round in circles of increasing radius and acquires high energy. The ion will finally come out of the dees in the direction indicated, through the window.

Energy of an ion. Let r_{max} be the radius of the outermost orbit described by the ion and v_{max} the maximum velocity gained by the ion in its final orbit. Then the equation for the motion of the ion in a magnetic field is

$$Bev_{\max} = \frac{mv_{\max}^2}{r_{\max}}$$

or $$v_{\max} = B\frac{e}{m}r_{\max} \qquad \text{...(5)}$$

∴ The energy of the ion

∴ $$E = \frac{1}{2}mv_{\max}^2 = \frac{B^2 r_{\max}^2}{2}\left[\frac{e^2}{m}\right] \qquad \text{...(6)}$$

The condition for acceleration of the ion in the inter-dee gap is that

$$\left.\begin{array}{l}\text{The time taken by the ion to}\\ \text{travel the semicircular path}\end{array}\right\} = \text{Half the time period of oscillation of the applied high frequency voltage}$$

i.e., $$\frac{\pi m}{Be} = \frac{T}{2} \text{ or } T = \frac{2\pi m}{Be}$$

∴ Frequency of the oscillator

$$f = \frac{Be}{2\pi m} \qquad \text{...(7)}$$

Hence the energy of the ion is given by

$$E = 2\pi^2 r_{\max}^2 f^2 m \qquad \text{...(8)}$$

The particles are ejected out of the cyclotron not continuously but as pulsed streams.

Limitations of the Cyclotron. The energies to which particles can be accelerated in a cyclotron are limited by the relativistic increase of mass with velocity. The mass of a particle, when moving with a velocity v is given by $m = \dfrac{m_0}{\sqrt{1-v^2/c^2}}$ where m_o is the rest mass and c the velocity of light.

According to equation (4),

$$\left.\begin{array}{l}\text{The time taken by the ion to}\\ \text{travel the semicircular path}\end{array}\right\} = t = \frac{\pi m}{Be} = \frac{T}{2}$$

$$\therefore \quad \text{Frequency of the ion}\} = n = \frac{1}{T} = \frac{Be}{2\pi m} \text{ or } n = \frac{Be\sqrt{1-v^2/c^2}}{2\pi m_0}.$$

Therefore, the frequency of rotation of the ion decreases with increase in velocity. The ions take longer time to describe their semicircular paths than the fixed period of the oscillating electric field. Thus, the ions lag behind the applied potential and finally they are not accelerated further. Due to this reason, the energy of the ions produced by the cyclotron is limited. This limitation can be overcome in the following two ways.

Now, the frequency of the ion $= Be\dfrac{\sqrt{1-v^2/c^2}}{2\pi m_0}$.

(*i*) Field variation. The frequency of the ion can be kept constant by increasing the magnetic field (B) at such a rate that the product $B\sqrt{1-v^2/c^2}$ remains constant. For this purpose, the value of the magnetic field B should increase, as velocity of the ion increases, so that the product $B\sqrt{1-v^2/c^2}$ remains unchanged. This type of machine in which the frequency of electric field is kept constant and magnetic field is varied is called *synchrotron.*

(*ii*) Frequency modulation. In another form of apparatus, the frequency of the applied A.C. is varied so that it is always equal to the frequency of rotation of the ion. This type of machine in which magnetic field is kept constant and the frequency of the applied electric field is varied is called a *frequency modulated cyclotron* or *synchro-cyclotron.*

EXAMPLE 1. *A cyclotron in which the flux density is 1.4 weber/m^2 is employed to accelerate protons. How rapidly should the electric field between the dees be reversed? Mass of the proton* $= 1.67 \times 10^{-27}$ *kg and charge* $= 1.6 \times 10^{-19}$ *C.*

SOL. Here, $B = 1.4$ weber/m^2; $m = 1.67 \times 10^{-27}$ kg; $e = 1.6 \times 10^{-19}$ C.

$$\therefore \qquad t = \frac{\pi m}{Be} = \frac{\pi(1.67\times10^{-27})}{1.4\times(1.6\times10^{-19})} = 2.342\times10^{-8}\ s.$$

EXAMPLE 2. *Deuterons in a cyclotron describe a circle of radius 0.32 m just before emerging from the dees. The frequency of the applied e.m.f. is 10 MHz. Find the flux density of the magnetic field and the velocity of deuterons emerging out of the cyclotron. Mass of deuterium* $= 3.32 \times 10^{-27}$ *kg;* $e = 1.6 \times 10^{-19}$ *C.*

SOL. We have, $\quad f = \dfrac{Be}{2\pi m} \quad \therefore\ B = \dfrac{2\pi mf}{e}$

Here, $\quad m = 3.32 \times 10^{-27}$ kg; $f = 10\ MHz = 10^7\ Hz$; $e = 1.6 \times 10^{-19}\ C$

$$\therefore \qquad B = \frac{2\pi(3.32\times10^{-27})10^7}{1.6\times10^{-19}} = 1.303\ \text{weber/m}^2$$

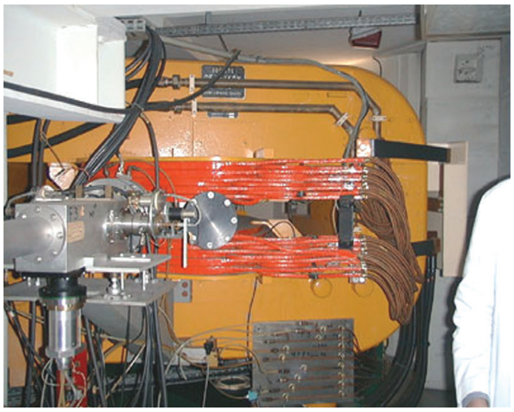
We have $\quad \dfrac{mv^2}{r_{\max}} = Bev \text{ or } v = \dfrac{Ber_{\max}}{m}$

Here, $B = 1.303$ weber/m^2 ; $e = 1.6 \times 10^{-19}$ C; $r_{\max} = 0.32\ m$ and $m = 3.32 \times 10^{-27}$ kg.

$$\therefore \qquad v = \frac{Be\,r_{\max}}{m} = \frac{1.303(1.6\times10^{-19})0.32}{3.32\times10^{-27}}$$

$$= 2.009 \times 10^{7}\ \text{ms}^{-1}.$$

30.5 The Synchrocyclotron



Synchrocyclotron

Synchrocyclotron is a modified form of the Lawrence cyclotron. This consists of only one dee placed in a vacuum chamber between the poles of an electromagnet (Fig. 30.5). Instead of the second dee, opposite the opening of the dee, there is a metal sheet connected to the earth. The alternating P.D. is applied between the dee and the metal plate (Fig. 30.6). The alternating potential applied to the "dee" is made to rise and fall periodically, instead of remaining constant. The frequency is changed at such a rate that as the ion lags a little due to the increase in mass caused by increase in velocity, the electric field frequency also automatically lags in variation. Hence the particle always enters the dee at the correct moment, when it can experience maximum acceleration. An advantage of using one dee is that it leaves sufficient space in the vacuum chamber for the ion source and the target. The pole-pieces of the magnet are of suitable shape such that the field decreases outwards from the centre. This ensures good focusing of the accelerated ions.

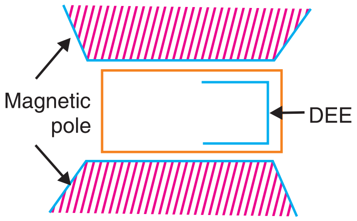


Fig. 30.5

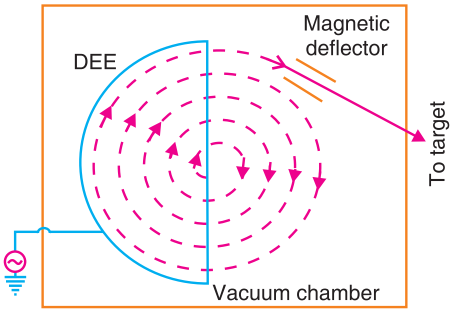
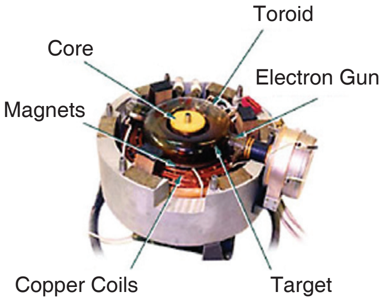


Fig. 30.6

30.6 The Betatron



Betatron

Betatron is a device to accelerate electrons (beta particles) to very high energies. It was constructed in 1941 by D.W. Kerst. The action of this device depends on the principle of a transformer.

Construction. It consists of a doughnut—shaped vacuum chamber placed between the pole-pieces of an electromagnet. The electromagnet is energised by an alternating current. The magnet produces a strong magnetic field in the doughnut. The electrons are produced by the electron gun (*FG*) and are allowed to move in a circular orbit of constant radius in the vacuum chamber

(Fig. 30.8). The magnetic field varies very slowly compared with the frequency of revolution of the electrons in the equilibrium orbit.

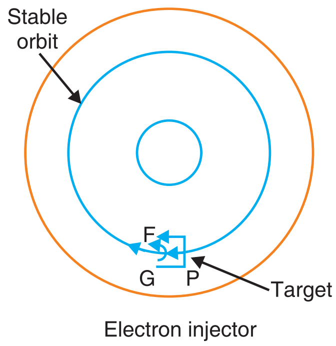


Fig. 30.7

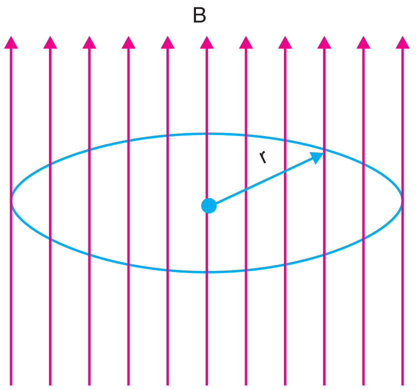


Fig. 30.8

The varying magnetic field, acting parallel to the axis of the vacuum tube, produces two effects on the electrons viz., (*i*) The changing flux due to the electromagnet produces the induced e.m.f. which is responsible for the acceleration of the electrons. (*ii*) The field of the magnet serves at the same time to bend the electrons in a circular path in the chamber and confine them to the region of the changing flux.

Theory. Consider the electron moving in an orbit of radius r (Fig. 30.9). Let ϕ be the flux linked with the orbit. The flux increases at the rate $d\phi/dt$ and the induced e.m.f. in the orbit is given by

$$E = -\frac{d\phi}{dt} \qquad \text{...(1)}$$

The work done on an electron of charge e in one revolution $= Ee = -e\dfrac{d\phi}{dt}$ ...(2)

Let F be the tangential electric force acting on the orbiting electron.

For one revolution, the path length is $2\pi r$. Then

the work done on the electron in one revolution = $F \times 2\pi r$

$$\therefore \qquad \mathrm{F} \times 2\pi r = -e\frac{d\phi}{dt}$$

or

$$F = -\frac{e}{2\pi r}\frac{d\phi}{dt} \qquad \text{...(3)}$$

When the velocity of the electron increases due to the above force, it will try to move into an orbit of larger radius. Because of the presence of the magnetic flux perpendicular to the plane of the electron orbit, the electron will experience a radial force inward given by

$$Bev = mv^2/r \qquad \text{...(4)}$$

Here B is the value of the magnetic field intensity at the electron orbit of constant radius r, v = velocity of the electron and m = mass of the electron. From (4),

the momentum of the electron = $mv = Ber$ ...(5)

From Newton's second law of motion,

$$F = \frac{d}{dt}(mv) = er\frac{dB}{dt} \qquad \text{...(6)}$$

To maintain the constant radius of the orbit, the values of F given in equations (3) and (6) must be numerically equal.

$$\therefore \qquad \frac{e}{2\pi r}\frac{d\phi}{dt} = er\frac{dB}{dt} \text{ or } d\phi = 2\pi r^2\ dB$$

Integrating, $\int_0^{\phi} d\phi = \int_0^{B} 2\pi r^2\ dB$

or $$\phi = 2\pi r^2 B \qquad ...(7)$$

If the uniform magnetic field B acts over an area πr^2, the magnetic flux $\phi' = \pi r^2 B$. Therefore the flux through the orbit is twice the flux enclosed by the orbit, if the magnetic field were to be uniform over the area. Equation (7) represents the condition under which a betatron works and is called *betatron condition.* This distribution of magnetic flux is obtained by the special pole-pieces where the magnetic field is greater at the centre of the orbit than at its circumference.

Fig. 30.9 shows the variation of magnetic field with time. Electrons are injected into the chamber when magnetic field just begins to rise. The electrons are then accelerated by the increasing magnetic flux linked with the electron orbit. During the time the magnetic field reaches its peak value, the electrons make several thousand revolutions and get accelerated. If they are allowed to revolve any more, the decreasing magnetic field would retard the electrons. Hence, the electrons are extracted at this stage by using an auxiliary magnetic field to deflect them from their normal course. The high energy electron beam can be made to strike the target, generating X-rays. Alternately the electrons can be made to emerge out of the apparatus and used for transmutation work.

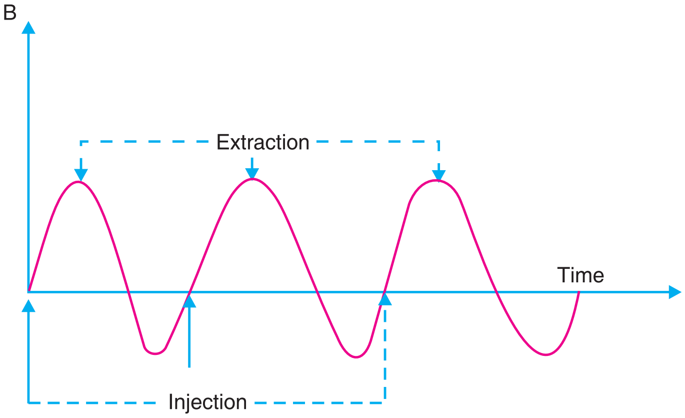


Fig. 30.9

EXAMPLE. *In a certain betatron the maximum magnetic field at orbit was 0.4 Wb/m*^{2}*, operating at 50 Hz with a stable orbit diameter of* 1.524 *m. Calculate the average energy gained per revolution and the final energy of the electrons.*

SOL. In the betatron, the electron velocities are nearly c.

$\therefore$ the total distance travelled in the acceleration time (*i.e.*, one quarter cycle) $= c\times\frac{T}{4} = c\times\frac{\pi}{2\omega}$

Total number of revolutions

$$= N = \frac{c\pi/2\omega}{2\pi r} = \frac{c}{4\omega r}.$$

Here, frequency $= f = 50$ Hz. $\therefore\ \omega = 2\pi f = 2\pi \times 50 = 100\ \pi$; $r = 0.762$ m, and $c = 3\times 10^8\ ms^{-1}$.

$$\therefore \qquad N = \frac{3\times10^8}{4(100\pi)0.762} = 3.132\times10^5$$

Let E be the final energy acquired by the electrons. Since the electrons must be treated relativistically,

momentum of the electron $= mv = E/c$

But $mv^2/r = Bev$ or $mv = Ber$ or $E = Berc$.

$$\therefore \qquad E = \frac{0.4(1.6\times10^{-19})(0.762)(3\times10^{8})}{1.6\times10^{-13}}MeV$$

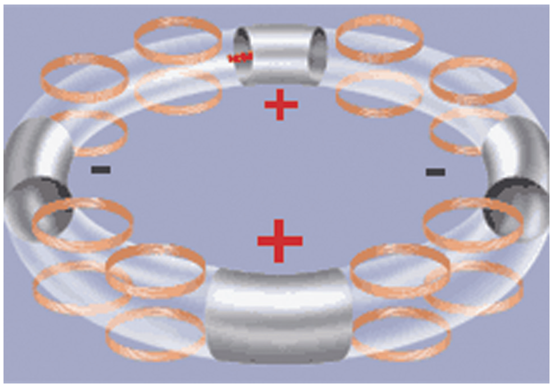
$$= 91.45\ MeV.$$

$$\text{Average energy gained per revolution} = \frac{91.45\times10^{6}}{3.132\times10^{5}} = 291.9\,eV\cdot$$

30.7 The Synchrotrons

There are two types of synchrotrons : (*a*) electron synchrotron and (*b*) proton synchrotron.

Electron synchrotron. The electron synchrotron is based on the principle of the combined working of betatron and cyclotron. Electrons are injected into an orbit of fixed radius at an initial energy of about 50 to 80 keV. The main accelerating tube, the torus, is made of glass or some plastic with a circular "dee" (*D*) made of a metal. An alternating potential is applied to the "dee" as shown in Fig. 30.10. A varying magnetic field is applied perpendicular to the torus. The radius of the orbit is kept constant by increasing the magnetic field as in a betatron. The increments of energy are given, as in a cyclotron, at the beginning and ending of the *D*. The electrons, after acceleration, are made to strike the required target. Using tungsten as target, very hard X-rays of energy about 300 MeV have been produced.



Synchrotrons.

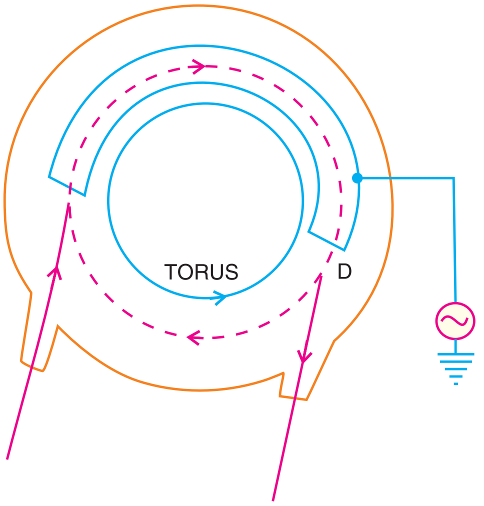


Fig. 30.10

30.8 The Proton Synchrotron (Bevatron, Cosmotron)

The proton synchrotron has been designed to provide protons of very high energy of the order of billion electron volts. The bevatron magnet consists of four segments connected by four equal straight sections (Fig. 30.11). The segments are quadrants of a circle of radius 30 feet and the connecting straight sections are 20 feet each. The protons are first accelerated by a linear accelerator and then injected into the synchrotron at one of the straight sections. The straight sections have an arrangement for applying the high frequency alternating potential. The magnetic field acts vertically upward. As the protons pass through one of the straight sections, they are accelerated and tend to have a bigger orbital radius. The magnetic field also increases at the same rate to constrain the proton beam to move in an orbit of constant radius. The output beam consists of a series of pulses of protons. By a suitable arrangement, the high energy particle may be made to strike a target.

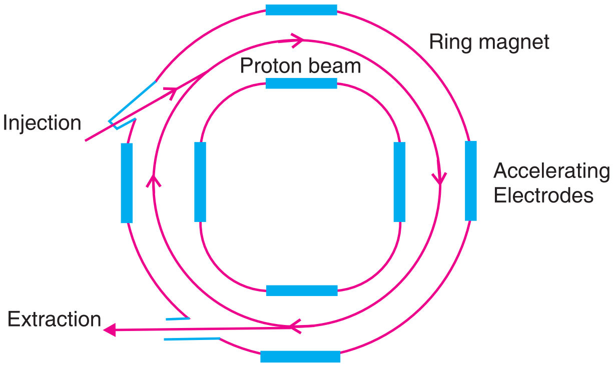


Fig. 30.11

Proton Synchrotron.

Theory. The frequency of revolution of the ion in a circular orbit is given by $n = Be/2\pi m$. In this derivation, we have neglected the relativistic variation in the mass of the accelerated ion. To take the relativistic effects into consideration, let us write

$$n = \frac{1}{2\pi}\frac{Bec^2}{mc^2} = \frac{1}{2\pi}\frac{Bec^2}{(m_0 c^2 + T)} \qquad \text{...(1)}$$

where m_0c^2 is the rest energy and T is the K.E. of the ion. Eqn. (1) gives the frequency of revolution of an ion in the absence of the straight sections. In the presence of four straight sections, each of length L, the frequency n' of circulation of protons is given by

$$n' = \frac{Bec^2}{2\pi(m_0c^2 + T)}\frac{2\pi r_0}{(2\pi r_0 + 4L)} \qquad \text{...(2)}$$

The proton is always guided by the magnetic field B to travel in the equilibrium orbit. Hence the following relation between B and the proton momentum p should hold at all times.

$$p = Ber_0 \qquad \text{...(3)}$$

If T is the K.E. of a proton of momentum p then

$$T + m_0c^2 = [(pc)^2 + (m_0c^2)^2]^{1/2} = [(Becr_0)^2 + (m_0c^2)^2]^{1/2}$$

or $$T^2 + 2T\,m_0c^2 + m_0{}^2c^4 = B^2\,e^2\,c^2\,r_0{}^2 + m_0{}^2c^4$$

$$\therefore \qquad B = \frac{[T\,(T + 2m_0c^2)]^{1/2}}{ecr_0} \qquad \text{...(4)}$$

Eqn. (4) describes the way in which the magnetic field B at the equilibrium orbit has to increase with the increase in the energy T of the proton as it circulates in the orbit. Substituting for B from Eqn. (4) in Eqn. (2), we get

$$n' = \frac{c[T\,(T + 2m_0c^2)]^{1/2}}{(m_0c^2 + T)(2\pi r_0 + 4L)} \qquad \text{...(5)}$$

In the proton synchrotron, the accelerating voltage has to be in phase with the circulation frequency of the proton in the equilibrium orbit. Hence the radio-frequency of the accelerating voltage should always match n' and vary in the manner in which n' varies with the proton energy T.

Example. *What radius is needed in proton synchrotron to attain particle energies of 10 GeV, assuming that a guide field of 1.8 Wb/m^2 is available?*

Sol. Equivalent mass of proton = 10 *GeV* + rest mass

$$= (10 + 0.938)\,GeV = 11.75\,u = 11.75 \times (1.66 \times 10^{-27})\text{ kg}$$

$$\text{Radius of the orbit} = \frac{mv}{Be} = \frac{(11.75 \times 1.66 \times 10^{-27})(3 \times 10^{8})}{1.8 \times (1.6 \times 10^{-19})} = 20.31\,m.$$

EXERCISE

1. Describe with diagram the construction and action of Van de Graaff generator. Mention its uses.

2. Can a Van de Graaff generator be used to accelerate electrons and protons?

3. Give the construction and working of a linear accelerator.

4. Describe the construction and action of a cyclotron. Discuss its limitations.

5. What is the difference between a cyclotron and a synchrocyclotron? Explain the theory and principle of working of a synchrocyclotron.

6. Describe giving necessary theory the working of a betatron. Show that to keep the electron in a constant radial motion, the magnetic induction in the area enclosed by the path has to be maintained at any instant equal to double the induction over the circular orbits.

7. Give an account of the theory, construction and working of a modern synchrotron. Give its uses.

8. A linear accelerator for the acceleration of protons to 45.3 MeV is designed so that, between any pair of accelerating gaps, the protons spend one complete radio-frequency cycle inside a drift tube. The r.f. frequency used is 200 Mc/sec.; (*i*) What is the length of the final drift tube? (*ii*) If the first drift tube is 5.35 cm long, at what K.E. are the protons injected into the linac? (*iii*) If the peak accelerating potential is 1.49×10^{6} volts, calculate the total length of the accelerator.

[**Ans.** (*i*) 0.4654 m (*ii*) 0.60 MeV (*iii*) 9.4 m]

9. Assume that in the 70 MeV betatron, the radius of the stable electron orbit is 28 cm. Calculate (*i*) the angular velocity of the electrons (*ii*) the frequency of the applied electric field and (*iii*) the value of the magnetic field intensity at the orbit, for this energy.

[**Ans.** (*i*) 1.07×10^{9} radian /sec. (*ii*) 1.7×10^{6} Hz (*iii*) 0.83 Wb m^{-2}]

10. Show that the radius of curvature R of the path of a particle inside the dees of a cyclotron is proportional to the $N^{1/2}$, where N is the number of times the particle has been accelerated across the space between the "dees".

11. Deuterons are accelerated in a fixed frequency cyclotron to a maximum dee orbit radius of 88 cm. The magnetic field is 1.4 weber/m^2. Calculate the energy of the emerging deuteron beam and the frequency of the dee voltage. What change in magnetic flux density is necessary if doubly charged helium ions are accelerated? Given atomic masses : $H^2 = 2.014102\,u$, $He^4 = 4.002603\,u$.

[**Ans.** 36.3 MeV ; 10.67 MHz; 0.01 Wbm^{-2}]

12. A fixed frequency cyclotron has an oscillator frequency of 12 MHz and dee radius of 0.55 m. It is used to accelerate deuterons. Calculate : (*a*) the magnetic flux density needed, and (*b*) energy to which deuterons are accelerated. [**Ans.** (*a*) 1.574 *Wb*/*m*^{2} (*b*) 17.95 *MeV*]

(c) If the above cyclotron is required to accelerate the doubly charged Helium ions, find how and by what amount magnet flux density will have to be changed at a fixed dee voltage.

Given : mass of helium ion = 6.642397×10^{-27} kg.

[**Hint.** Let B_1 and B_2 be the values of the magnetic flux densities when the cyclotron accelerates two ions having charges Z_1e and Z_2e and masses m_1 and m_2 respectively. Then,

$$B_1 = 2\pi m_1 f/Z_1 e \text{ and } B_2 = 2\pi m_2\, f/Z_2 e$$

$$\therefore \qquad B_1 - B_2 = \frac{2\pi f}{e}\left(\frac{m_1}{Z_1} - \frac{m_2}{Z_2}\right)$$

$$= \frac{2\pi(12\times10^{6})}{1.6\times10^{-19}}\left(\frac{3.34245\times10^{-27}}{1} - \frac{6.642397\times10^{-27}}{2}\right).$$

$$= 0.01008 \text{ Wb m}^{-2} \text{ (to be decreased)}].$$

13. A betatron has the following parameters : Magnet current supply frequency = 50 *Hz*. Peak magnetic flux density at the orbit = 0.4 Wb m^{-2}. Electron orbit radius = 0.75 m. Calculate (*a*) the final energy of the electrons (*b*) energy gained per revolution and (*c*) total time of flight of the electrons.

[**Ans.** (*a*) 90.01 *MeV* (*b*) 282.7 *eV* (*c*) 4.999×10^{-3} seconds]

14. 0.5 *MeV* protons are injected into a 50 *MeV* linear accelerator powered by a 200 *MHz* radio-frequency supply. Find the approximate lengths of the first and the last drift tubes.

[**Ans.** 0.02446 m; 0.2354 m]

15. Calculate the average energy per revolution and final energy gained by electrons in a betatron to which is applied a maximum magnetic field of 0.5 tesla operating at 60 Hz in a stable orbit of diameter 2 metres. [**Ans.** 750 eV, 150 MeV].

16. What total accelerator voltage will impart protons a velocity of 95% of the speed of light? (Rest mass of proton = $m_0 = 1.672 \times 10^{-27}$ kg).

[**Hint.**

$$m = \frac{m_0}{\sqrt{1-\frac{v^2}{c^2}}} = \frac{1.672\times10^{-27}}{\sqrt{1-\left(\frac{95}{100}\right)^2}} = 5.356\times10^{-27}\text{ kg}$$

$$\therefore \quad \Delta m = (m - m_0) = (5.356 - 1.672) \times 10^{-27}\text{ kg}$$

$$\therefore \quad \text{KE} = \Delta m.c^2 = (3.684 \times 10^{-27}) \times (3 \times 10^8)^2\text{ J} = 2.07 \times 10^9\text{ eV}$$

Hence accelerating voltage = 2.07 BeV].

17. In a synchrocyclotron, the magnetic flux density decreases from 1.5 Wb m^{-2} at the centre of the magnet to 1.43 Wb m^{-2} at the limiting radius of 2.06 m. Calculate the range of frequency modulation required for deuteron acceleration and the maximum kinetic energy of the deuterons. The rest mass of the deuteron is 3.34×10^{-27} kg; the electronic charge is 1.6×10^{-19} *C*.

[**Hint**.

$$\frac{m_0 v}{\sqrt{1-(v^2/c^2)}} = Ber \qquad \text{...(1)}$$

where *B* is the magnetic induction at the limiting radius *r* and *v* is the final velocity attained by the deuteron.

Here,

$$m_0 = 3.34 \times 10^{-27}\ kg\ ;\ B = 1.43\ Wb\ m^{-2}\ ;$$

$$e = 1.6 \times 10^{-19}\ C\ ;\ r = 2.06\ m,\ c = 3 \times 10^8\ ms^{-1};\ \therefore\ v = 1.277 \times 10^8\ ms^{-1}$$

$$\text{Frequency of revolution} = \frac{v}{2\pi r} = \frac{1.277\times10^8}{2\pi\times2.06}\,Hz = 9.87\,MHz.$$

The initial frequency of the oscillator must equal the simple cyclotron frequency

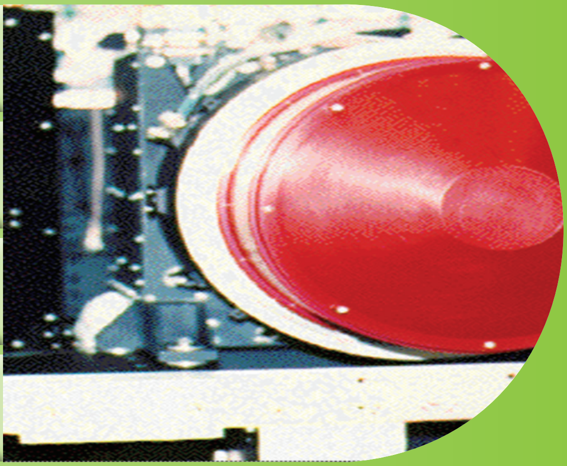
$$f = \frac{Be}{2\pi m} = \frac{1.5\times(1.6\times10^{-19})}{2\pi\times3.34\times10^{-27}} = 11.44\,MHz.$$

Therefore the range of frequency modulation is from 11.44 *MHz* to 9.87 *MHz*.

$$E = m_0 c^2\left[\frac{1}{(1-v^2/c^2)^{12}} - 1\right] = 197.2\,MeV].$$

CHAPTER

31



RADIOACTIVITY

AT A GLANCE

31.1	Discovery of Radioactivity	**31.2**	Natural Radioactivity
31.3	Alpha, Beta and Gamma Rays	**31.4**	Properties of Alpha Rays
31.5	Properties of Beta Rays	**31.6**	Properties of Gamma Rays
31.7	Determination of E/M of Alpha Particles		
31.8	Determination of the Charge of Alpha Particles		
31.9	Velocity of Alpha Particles	**31.10**	Range of α-particles
31.11	Experimental Measurement of the Range of α-particles		
31.12	Alpha Particle Disintegration Energy	**31.13**	Alpha Particle Spectra
31.14	Theory of Alpha Decay	**31.15**	Gamow's Theory of Alpha Decay
31.16	The Nature of Beta Particles		
31.17	Determination of e/m of Beta Particles: Kaufmann's Experiment		
31.18	Bucherer's Experiment-Increase of Beta Particle Mass with Velocity		
31.19	Beta Ray Spectra	**31.20**	Magnetic Spectrograph
31.21	Origin of the Line and Continuous Spectrum		
31.22	The Neutrino Theory of Beta Decay	**31.23**	Gamma Ray spectra
31.24	Determination of the Wavelength of Gamma Rays		
31.25	Origin of γ-rays	**31.26**	Nuclear Isomerism
31.27	Internal Conversion		
31.28	Mossbauer Effect		
31.29	Soddy Fajan's Displacement Law	**31.30**	Law of Radioactive Disintegration
31.31	The Mean Life	**31.32**	Measurement of Decay Constants
31.33	Units of Radioactivity	**31.34**	Law of Successive Disintegration
31.35	Radioactive Dating: The Age of the Earth		
31.36	Biological Effects of Nuclear Radiations		
31.37	Non-Conservation of Parity in Beta Decay		

31.1 Discovery of Radioactivity

Radioactivity was discovered by Henri Becqerel in 1896. He found that a uranium salt wrapped up in paper emitted certain penetrating radiations which affected a photographic plate. He found that these radiations were given off spontaneously without any need to any initial excitation of the uranium salt. He showed that the radiation is not influenced by any external agency. Further investigations by Madame Curie, Piere Curie, Rutherford etc., showed that the phenomenon was exhibited by heavy elements like uranium, polonium, radium, thorium etc. Radioactivity involves the spontaneous transmutation of one element into another.

Henri Becqerel
(1852-1908)

31.2 Natural Radioactivity

The phenomenon of spontaneous emission of highly penetrating radiations from heavy elements of atomic weights greater than about 206, occuring in nature, is called *natural radioactivity.* The elements which exhibit this property are called *radioactive elements.* The atoms of radioactive elements emit radiations composed of three distinct kinds of rays (α, β and γ). In the process, the elements break up, leading to an irreversible self-disintegration. The activity is *spontaneous.* Radioactivity is unaffected by any external agent like high temperature, high pressure, large electric and magnetic fields etc.

The modern technique of artificial transmutation of elements has been able to produce radioactivity in many other elements, much lighter than those that occur in nature. Radioactivity induced in an element by bombarding it with α-particles, neutrons, protons, and other particles or radiations is called artificial radioactivity. Hence there is a distinction between *natural radioactivity i,e., of elements* as found in nature and *induced* radioactivity, *i.e.*, artificially produced. But whatever be its origin, the activity is always spontaneous. The artificial radioactive elements usually, but not always, have short life-times. They emit electrons, positrons, and other particles, as well as γ-rays in their decomposition.

31.3 Alpha, Beta and Gamma Rays

The existence of three distinct types of radiation is demonstrated by the following simple experiment. A small quantity of radium (*R*) is placed at the bottom of a small hole drilled in a lead block [Fig. 31.1] A fairly parallel beam of radiation from *R* will issue through the hole. This lead block is placed inside an evacuated chamber to avoid absorption of the rays. A photographic plate (*P*) is placed at a short distance above the lead block to receive the rays. A strong magnetic field is applied at right angles to the plane of the figure and directed away from the reader. After a fairly long exposure, the photographic plate is developed. Three distinct lines will be found on the photographic plate. The α–particles will be deflected towards the left, indicating that they are *positively* charged. The β-particles will be deflected towards the right, showing that they are *negatively* charged. The γ-rays are not deflected and hit the plate *P* straight. This shows that the γ-rays are uncharged or neutral rays. If an electrostatic field is applied, in place of the magnetic field (Fig. 31.2), the β-rays are deflected towards the positive plate, the α-rays towards the negative and γ-rays do not bend at all.

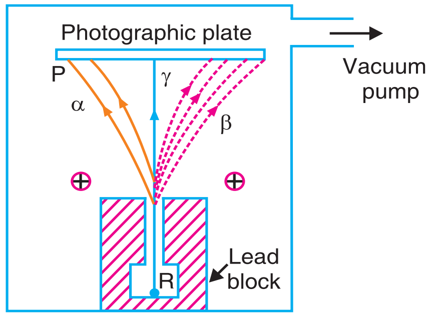


Fig. 31.1

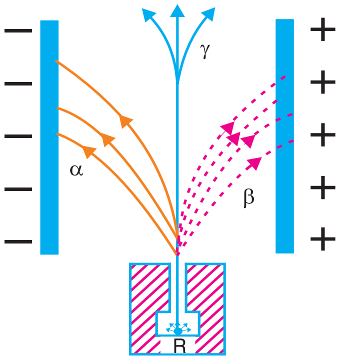


Fig. 31.2

31.4 Properties of Alpha Rays

(1) An α-particle is a helium nucleus consisting of two protons and two neutrons. It carries two units of positive charge.

(2) α-particles shot out from radioactive substances have high velocities, ranging from 1.4×10^7 to 1.7×10^7 m/s. They move along straight lines. Their tracks can be observed in the Wilson's cloud chamber.

(3) They produce intense ionisation in the gas through which they pass. Their ionising power is 100 times greater than that of β-rays and 10,000 times greater than that of γ-rays.

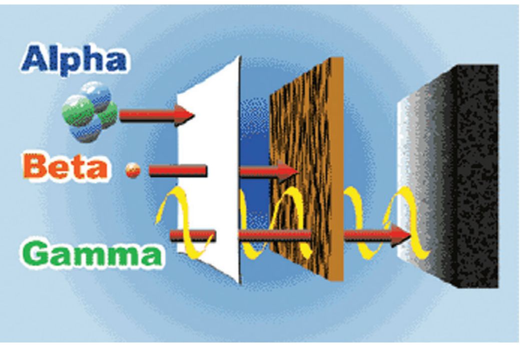
(4) They affect a photographic plate but the effect is very feeble.

(5) They produce fluorescence when they fall on substances like barium platinocyanide or zinc sulphide. The fluorescence can be observed through a low power microscope.

(6) They are deflected by electric and magnetic fields. This shows that they are charged particles.

(7) They are scattered by nuclei of heavy elements like gold.

(8) They produce a heating effect. The evolution of heat is due to the stoppage of α, β and γ-rays by the radioactive substance.



Properties of α, β, γ-Rays.

(9) The value of E/M for α-particles was found by Rutherford to be half of that for hydrogen ion. The charge on each α-particle (E) was found to be twice that of a hydrogen ion. Thus the mass of the α-particles was shown to be four times that of hydrogen. Therefore, an α-particle is the nucleus of the helium atom.

31.5 Properties of Beta Rays

(1) β-particles possess negative charge and mass equal to that of an electron.

(2) All the β-particles emitted from a substance do not have the same velocity. The velocities of the β-particles emitted by a substance range continuously from $0.3c$ to $0.99c$. At high velocities, e/m is found to decrease indicating an increase in mass of the particles, according to Einstein's equation,

$$m = \frac{m_0}{\sqrt{1-v^2/c^2}}.$$

(3) Their ionising power is low and hence, range is large.

(4) They affect a photographic plate.

(5) They produce fluorescence in barium platinocyanide, willemite etc.

(6) They are deflected by electric and magnetic fields. Their direction of deflection indicates that they are negatively charged particles.

(7) They penetrate through thin metal foils and their penetrating power is greater than that of α-rays.

(8) The value of e/m for β-rays was proved to be the same as that of cathode rays. The charge on a β-particle was observed to be the same as the charge on an electron. Therefore, β-rays are identical with electrons.

31.6 Properties of Gamma Rays

(1) γ-rays are electromagnetic waves of very short wavelength. The wavelengths of γ-rays range from 0.005 Å to 0.5 Å. They are not charged particles. They travel with the velocity of light.

(2) They produce fluorescence and affect a photographic plate.

(3) They ionise the gas through which they pass but the ionisation produced is very small.

(4) They are more penetrating than even β-rays. They can even pass through an iron plate of 30 cm thickness.

(5) They are not affected by electric and magnetic fields.

(6) They are diffracted by crystals like X-rays.

THE ALPHA RAYS

31.7 Determination of E/M of Alpha Particles

The experiment of Rutherford and Robinson to measure the specific charge of α-rays is described below.

Principle. A beam of α-particles is subjected to electric and magnetic fields, and from the deflections produced, the value of E/M is determined.

The apparatus used for the magnetic deflection experiment is shown in Fig. 31.3. A fine wire coated with radium (*Ra C*) which emits α-particles is placed in a cavity L of a lead block. The α-particles emanating from the source are limited by a slit S and are incident on a photographic plate P. The whole arrangement is kept in an evacuated vessel V. A uniform magnetic field of flux density B is applied perpendicular to the plane of the paper. Under the action of the magnetic field, the α- particle will describe an arc of a circle of radius r. On reversing the magnetic field, the path of the particle is reversed. The α-particles reach the points P_1 and P_2 respectively before and after reversing the magnetic field. The velocity v of the α-particles is more or less the same for all the

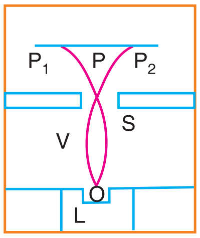


Fig. 31.3

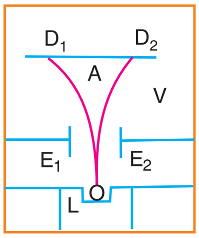


Fig. 31.4

particles in the case of the *Ra C* source. If *E* and *M* respectively be the charge and mass of the α-particle, then

$$BEv = \frac{Mv^2}{r}$$

or
$$\frac{E}{Mv} = \frac{1}{Br} \qquad \text{...(1)}$$

The radius *r* of the circular path can be easily measured, since the path is defined by the three points LSP_1 or LSP_2. Since *B* and *r* are known, *E*/*M* can be found if *v* is determined.

To determine E/Mv^2, an apparatus as in Fig. 31.4 is used. In this experiment, α-particles from the same source are subjected to the action of an electric field. E_1 and E_2 are two metallic plates. A strong electric field is maintained between the plates. The α-particles passing through the electric field are deflected at right angles to their path. After travelling a parabolic path in the electric field, the particles are recorded on a photographic plate at D_1 and the deflection is AD_1. When the field is reversed, the trace is obtained at D_2 and the deflection is AD_2. The mean deflection *d* is found.

Let l = the length over which the electric field acts.
L = distance of photographic plate from the field,
X = intensity of the electric field and
d = deflection on the photographic plate.

Then
$$d = \frac{XE}{M}\frac{l}{v}\frac{L}{v}$$

or
$$\frac{E}{Mv^2} = \frac{d}{XlL} \qquad \text{...(2)}$$

Knowing $\frac{E}{Mv}$ from Eqn. (1) and $\frac{E}{Mv^2}$ from Eqn. (2), we can get the values of both *E*/*M* and *v*. The value for *E*/*M* of α-particles is 4.82×10^7 C/kg. This is very close to the value of *E*/*M* calculated for doubly ionised helium.

31.8 Determination of the Charge of Alpha Particles

A known quantity of *Ra C* is placed in a shallow dish *R* (Fig. 31.5). The dish is covered with aluminium foil which absorbs the atoms from which α-particles are emitted. These atoms cannot escape because of recoil. The α-particles, after passing through a window (*W*) of known area covered with a very thin aluminium foil, fall upon the collecting plate *P*. *P* is connected to an electrometer. The apparatus is kept exhausted. The charge received by the plate *P* in a given time *t* is found from the electrometer reading. A strong magnetic field prevents the β-rays emitted by the source from reaching the collecting plate. The number of α-particles falling on the plate is determined by using the charge carried by each α-particle is estimated.

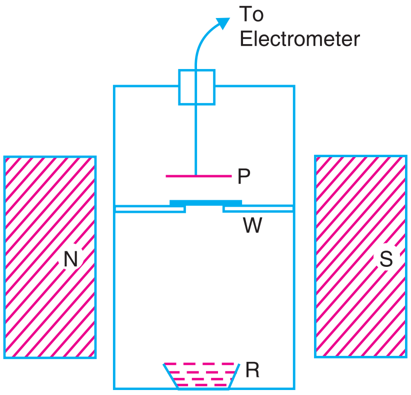


Fig. 31.5

$$E = 3.19 \times 10^{-19}\ C.$$

Identification of α-particles

Charge on the α-particle = $E = 3.19 \times 10^{-19}\ C$.

Electronic charge = $e = 1.6 \times 10^{-19}\ C$.

Hence $E = 2e$.

$$\text{Mass of } \alpha\text{-particles } M = \frac{E}{(E/M)} = \frac{3.19\times10^{-19}}{4.82\times10^{7}} = 6.62\times10^{-27}\ \text{kg}\cdot$$

Comparing this with the mass of the hydrogen atom

$$\frac{M}{M_H} = \frac{6.62\times10^{-27}}{1.67\times10^{-27}} \approx 4 \quad \text{or} \quad M = 4M_H.$$

Thus the α-particle has the mass of a helium atom and carries two positive elementary charges. This shows that the α-particle is a helium atom that has lost both its electrons, *i.e.*, a bare helium nucleus.

31.9 Velocity of Alpha Particles

The velocity of emission of α-particles is characteristic of the isotope from which they emanate. Rutherford's experiments gave the velocity of the alpha-particles from *Ra C* as 2×10^7 ms^{-1}. In general, the velocity of α-particles varies from 1.45 to 2.2×10^7 ms^{-1} with different radioactive sources.

31.10 Range of α-particles

α-particles are capable of ionising a gas. In the course of its passage through a gas, the α-particle gets slowed down by losing energy in ionising the gas particles, until its energy falls below the ionisation potential of the gas. Then the α-particle captures two electrons and becomes a neutral helium atom. The distance, the α-particle travels in the gas in this manner, is called its range. Beyond this distance (range), the three properties of the α-rays (ionising power, photographic action and fluorescence effect) disappear simultaneously. The range is usually expressed as so many cm in air at 0.76 m pressure and 15°C. The range depends on (*i*) the initial energy of the α-particles (*ii*) the ionisation potential of the gas and (*iii*) the chances of collision between the α-particle and the gas particles *i.e.*, on the nature and pressure of the gas.

Definition : *The distance through which an α-particle travels in a substance before coming to rest is called the 'range' of the particle in that substance.*

Electronic Setup for the Range of α-Particles in Air.

31.11 Experimental Measurement of the Range of α-particles

(*i*) Fig. 31.6 shows Bragg's apparatus for determining the range of α-particles in a gas at different pressures. *T* is an evacuated glass tube. The radioactive substance *R* (in a lead block) is placed on the platform of the adjustable rod *C*. This rod can be raised up or down. The tube can be filled with the gas at any desired pressure which is measured by a manometer. The plate *P* is connected to the electrometer. The ionisation produced by the particles emitted by the source is measured by an ionisation chamber. The ionisation chamber consists of the plate *P* and a grid *G*. Between these electrodes is maintained a *P.D.* which ensures that an electrometer connected to the chamber measures the saturation ionisation current in the chamber. The second grid *G'* prevents the diffusion of ions from the chamber.

The tube is filled with the required gas. The rod *C* is adjusted and the α-particles reaching the grid *G* pass into the space between the grid *G* and plate *P*. The particular position of the platform is found when the ionisation current falls to zero. This is detected by the electrometer. The distance between the source and the grid *G* gives the range of the α-particles for the particular element at a given temperature and pressure in a particular gas.

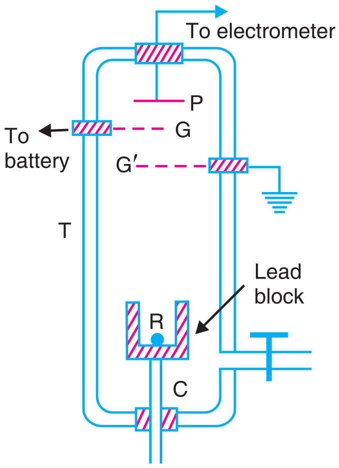


Fig. 31.6

It is experimentally established that the range is inversely proportional to the density of the gas, *i.e.*, the range is inversely proportional to the pressure of the gas. *Specific ionisation is the number of ions produced by* α-*particles per unit length of the path.* Bragg found that the curve relating the specific ionisation with the distance from the source was as shown in Fig. 31.7. It is seen that the graph shows the so called "Bragg hump" near the end of the range and also a slight "tail" or "*straggling effect*". The increase of ionisation near the end of the range can be understood because at this point the α-particles have relatively low velocities and, therefore, have a greater probability of colliding with the gas atoms. The "tail" effect arises because all the α-particles are not emitted from the source with the same initial velocity. All of them do not, therefore, reach the end of their range simultaneously.

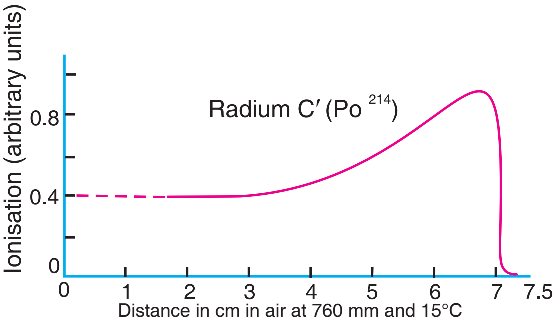


Fig. 31.7

(*ii*) **Geiger and Nuttal experiment.** The α-particle source is in the form of a thin film on a small metal disc *M* (Fig. 31.8). It is mounted at the centre of a spherical glass bulb *B*. The bulb is coated inside with silver and a high *P.D.* is applied between the silver film and the disc *M*. The saturation ionisation current in the bulb produced by the passage of the α-particles through the gas in the bulb is measured by an electrometer.

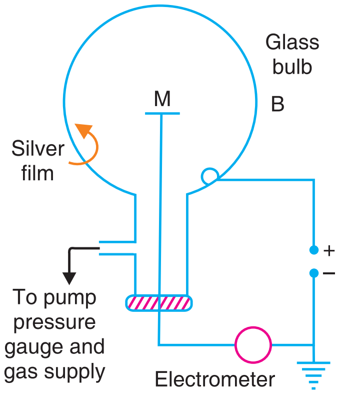


Fig. 31.8

The gas under study can be admitted at any desired pressure into the bulb. The saturation current for different gas pressures is measured. A graph is plotted between the ionisation current and the corresponding pressure *P* of the gas. We

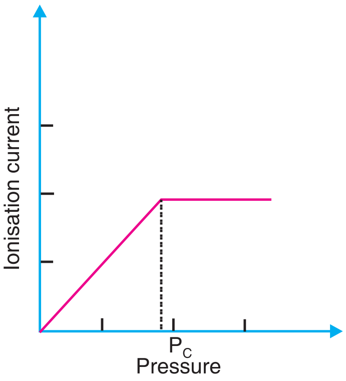


Fig. 31.9

get curves of the type shown in Fig. 31.9. The ionisation at first increases directly with the pressure. At a certain critical pressure P_c, the density of the gas is such that the particles are just stopped on the inner surface of the bulb, having produced the maximum ionisation possible, before reaching the bulb. If the pressure is increased beyond P_c, no rise in the current is observed, since the α-particles get absorbed in a smaller path and do not reach the wall. The radius of the bulb thus gives the range of the α-particles at the gas pressure P_c. As the range is inversely proportional to pressure, the range at normal pressure can be calculated.

Geiger's law. Geiger studied the relation between the range (R) of an α-particle and its velocity of emission (v). He found that the range R is proportional to the cube of the velocity. $R \propto v^3$ or $R = av^3$, where a is a constant.

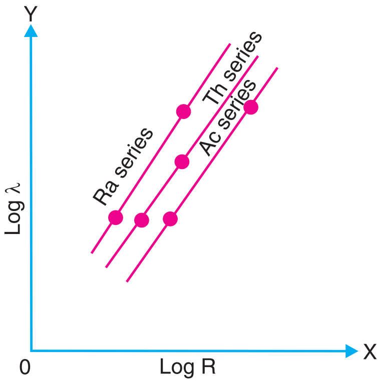


Fig. 31.10

Range-energy relation. The energy E of the emitted particle is directly proportional to the square of the velocity v. Hence Geiger's law can be expressed in the form $R = bE^{3/2}$ where b is a constant.

Geiger-Nuttal law. The range R of an α-particle and the disintegration constant λ of the radioactive element that emits it are related as follows:

$$\log \lambda = A + B \log R.$$

This relation is called Geiger-Nuttal law. If log λ are plotted against log R for the different α-emitters in the three series, three nearly parallel straight lines are obtained, one for each series (Fig. 31.10). In the relation $\log \lambda = A + B \log R$, the constant B is the same for all the series while A is different for the different series. According to this relation, when the disintegration constant is high, the range is also high. Since the range also depends on the energy, we conclude that radioactive substances of large decay constants emit high energy α-particles.

This law is helpful in determining roughly the decay constants of radioactive substances of very short or very long lives. Experimentally measuring the ranges of α-particles of such radioelements, the respective decay constants can be obtained by extrapolation from the curves representing the Geiger-Nuttal relation.

31.12 Alpha Particle Disintegration Energy

When an α-particle is emitted from a nucleus, the nucleus recoils in order to conserve momentum. Let m and v be the mass and velocity of the α-particle. Let M and V be the mass and velocity of the daughter nucleus. According to the law of conservation of momentum,

$$mv = MV.$$

The sum of the kinetic energies of the α-particle and the product nucleus is called alpha disintegration energy (E).

$$E = \frac{1}{2}mv^2 + \frac{1}{2}MV^2 = \frac{1}{2}mv^2 + \frac{1}{2}M\left(\frac{mv}{M}\right)^2$$

$$\therefore \quad E = \frac{1}{2}mv^2\left(1 + \frac{m}{M}\right).$$

The α disintegration energy can thus be easily obtained by multiplying the K.E. of the α-particles by the factor $(1 + m/M)$.

Example. *Polonium-212 emits α-particles whose K.E. is* 10.54 *MeV. Determine the* α *disintegration energy.*

Sol. When polonium - 212 emits an α-particle, the mass of the daughter nucleus is 208 units. Here, $M = 208$ units, $m = 4$ units;

$$\text{K.E. of the α-particles} = \frac{1}{2}mv^2 = 10.54\,MeV.\,E = ?$$

$$E = \frac{1}{2}mv^2\left(1 + \frac{m}{M}\right) = 10.54\left(1 + \frac{4}{208}\right) = 10.74\,MeV$$

31.13 Alpha Particle Spectra

Setup for taking Alpha Particle Spectra.

It was thought earlier that the nucleus of a radioactive substance emitted alpha particles of some fixed energy. By measuring the energies of α-particles accurately using a magnetic spectrograph, it was shown that several radioactive nuclei emitted α-particles of different energies falling within a narrow range. As an example, consider a *ThC* nucleus which decays into a *ThC"* ($_{81}Tl^{208}$) nucleus after emitting an α-particle. *ThC* emits five groups of α-particles with energies 5.709, 5.728, 5.873, 6.161 and 6.201 MeV. The emission of groups of α-particles with different energies suggests that nuclei may exist in a number of discrete, excited energy states above the ground state. The parent nucleus decays from its ground state to anyone of the several excited states or to the ground state of the daughter nucleus. The daughter nucleus *Th C"* formed as a result of disintegration of *ThC* exists in several excited energy states *a*, *b*, *c*, *d*, *e*. *e* is the ground state and *a* is the highest excited state of the daughter nucleus. Each excited state has got some energy which it gives up in the form of γ radiation during its transition to the ground state. When a parent nucleus goes from its ground state to the ground state of the daughter nucleus, it emits during such a transition, an α-particle of maximum energy (α_1 in the energy diagram, Fig. 31.11). Any transition from the ground state of the parent nucleus to any one of the excited states of the daughter nucleus will give rise to α-particle with smaller energy (α_2, α_3, α_4 and α_5 in Fig. 31.11). The transition from any excited state to the ground state involves the emission of γ-rays. The energy of the γ-ray photons have been measured and found to be equal to the energy difference between one of the excited energy states and the ground state. Thus, the α-spectrum provides a conclusive evidence of the existence of *discrete energy levels in a nucleus.*

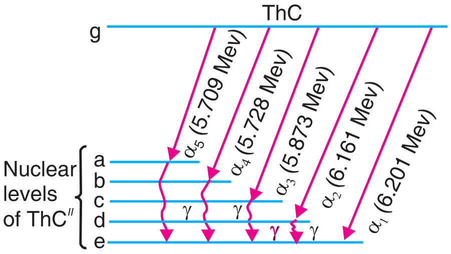


Fig. 31.11

Decay of parent nucleus ThC into various excited levels of the daughter nucleus ThC"

31.14 Theory of Alpha Decay

Only heavy nuclei with $A > 200$ undergo α-decay. The α-particles emitted from nuclei have a discrete energy spectrum and consist of several groups. Usually the most intensive is the group with α-particles of highest energy. The existence of several groups of α-particles is called the *fine structure of the* α*-spectrum.*

Before emission, the α-particle can be considered to be inside the nucleus. Coulomb's law is applicable when the α-particle is outside the nucleus. When the α-particle is inside the nucleus or very close to it, Coulomb's law does not hold good. For an α-particle of charge $2e$ and a nucleus of charge $(Z–2)\,e$ separated by a distance r, the potential energy is given by

$$V_r = \frac{2e(Z-2)e}{4\pi\varepsilon_0 r} = \frac{2(Z-2)e^2}{4\pi\varepsilon_0 r}.$$

Here $(Z - 2)$ is the atomic number of the daughter nucleus. An attractive potential, called the potential well, represents the position (Fig. 31.12). When $r < r_0$, the α-particle is within the potential well and it is bound by the nucleus. Here r_0 is equal to the sum of the radii of the nucleus and the α-particle. According to wave mechanics, an α-particle can have different energy levels E_0, E_1, E_2, etc., within the potential well. Only an α-particle at a level E_3 or above can come out of the potential well. In the case of radioactive elements, the height of the potential barrier is about 9 MeV. But uranium emits α-particles of energy 4 MeV. Classical mechanics cannot explain how a particle having an energy of 4 MeV can come out of a well having a potential barrier of 9 MeV.

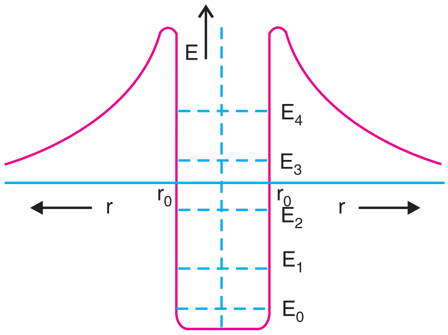


Fig. 31.12

The escape of an α-particle from a radioactive nucleus can be explained on the basis of wave mechanics and Schrodinger equation. According to it, it is possible for an α-particle to *leak* through the potential barrier, even though its K.E. is less than the potential energy of the height of potential barrier. This probability of leaking of an α-particle through the barrier is called *tunnel effect*. The α-particle within the nucleus must present itself again and again at the barrier surface until conditions are ripe for penetration or leakage.

31.15 Gamow's Theory of Alpha Decay

Classical physics fails to explain α-decay. Quantum mechanics provides a successful explanation of the problem of α-decay. The basic notions of this theory are :

(1) An alpha particle may exist as an entity within a heavy nucleus.

(2) Alpha particle is in constant motion and bounces back and forth from the barrier walls. In each collision with 'wall' there is a definite probability that the particle will leak through the potential barrier. Let ν be the frequency with which the α-particle collides with the walls in order to escape from the nucleus and P the probability of transmission in each collision. Then the decay probability per unit time (*i.e.*, disintegration constant λ) is given by

$$\lambda = \nu P \qquad ...(1)$$

Suppose that at any moment only one α-particle exists as such in a nucleus and that it moves back and forth along a nuclear diameter.

Then $$v = \frac{v}{2R} \qquad \text{...(2)}$$

where v is the alpha particle velocity and R is the nuclear radius.

Since $V > T$ [where V = height of the potential barrier and T = K.E. of α-particle], according to classical physics $P = 0$ [Fig. 31.13]. But in quantum mechanics a moving particle is regarded as a wave, and the result is a small but definite value for P. The probability P can be calculated quantum mechanically using WKB perturbation theory. According to this theory,

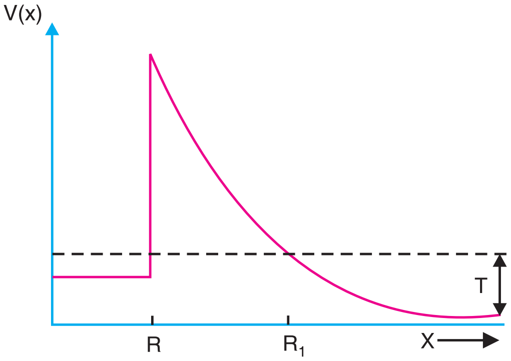


Fig. 31.13

$$\log_e P = \frac{-2}{(h/2\pi)} \int_R^{R_1} \sqrt{2m\{V(x) - T\}}\, dx \qquad \text{...(3)}$$

where m = mass of the α-particle.

$$V(x) = \frac{2Ze^2}{4\pi\varepsilon_0 x}$$

is the eletrostatic P.E. of an alpha particle at a distance x from the centre of a nucleus of charge Ze (Ze is the charge of the daughter nucleus *i.e.*, nuclear charge minus the alpha particle charge of $2e$)

R = Nuclear radius

T = K.E. of α-particle such that $T < V(x)$.

The region from $x = R$ to $x = R_1$ is called the thickness of the barrier.

We therefore have:

$$\log_e P = \frac{-2}{(h/2\pi)} \int_R^{R_1} \sqrt{2m\left(\frac{2Ze^2}{4\pi\varepsilon_0 x} - T\right)}\, dx \qquad \text{...(4)}$$

When $x = R_1$, $T = V$. $\therefore T = \dfrac{2Ze^2}{4\pi\varepsilon_0 R_1}$

$$\therefore \quad \log_e P = \frac{-2}{(h/2\pi)} (2mT)^{1/2} \int_R^{R_1} \left(\frac{R_1}{x} - 1\right)^{1/2} dx$$

$$\log_e P = \frac{-2}{(h/2\pi)} (2mT)^{1/2} R_1 \left[\cos^{-1}\left(\frac{R}{R_1}\right)^{1/2} - \left(\frac{R}{R_1}\right)^{1/2}\left(1 - \frac{R}{R_1}\right)^{1/2}\right] \qquad \text{...(5)}$$

The width of the potential barrier is very large compared with the nuclear radius *i.e.*, $R_1 >> R$. Therefore

$$\cos^{-1}\left(\frac{R}{R_1}\right)^{1/2} \approx \frac{\pi}{2} - \left(\frac{R}{R_1}\right)^{1/2} \text{ and } \left(1 - \frac{R}{R_1}\right)^{1/2} \approx 1$$

Hence $$\log_e P = \frac{-2}{(h/2\pi)}(2mT)^{1/2} R_1 \left(\frac{\pi}{2} - 2\left(\frac{R}{R_1}\right)^{1/2}\right)$$

Substituting $R_1 = 2\,Ze^2 / 4\pi\,\varepsilon_0\,T$, we have

$$\log_e P = \frac{4e}{(h/2\pi)}\left(\frac{m}{\pi\varepsilon_0}\right)^{1/2} Z^{1/2} R^{1/2} - \frac{e^2}{(h/2\pi)\varepsilon_0}\left(\frac{m}{2}\right)^{\frac{1}{2}} Z\,T^{-1/2} \quad ...(6)$$

or $$\log_e P = 2.97\,Z^{1/2} R^{1/2} - 3.95\,ZT^{-1/2}$$

where R is in fermi, and T is in MeV.

Since $$\lambda = \nu P, \log_e \lambda = \log_e \nu + \log_e P = \log_e\left(\frac{v}{2R}\right) + \log_e P$$

$\therefore$ $$\log_e \lambda = \log_e\left(\frac{v}{2R}\right) + 2.97\,Z^{1/2} R^{1/2} - 3.95\,Z\,T^{-1/2} \quad ...(7)$$

Changing the base of log (from e to 10),

$$\log_{10} \lambda = \log_{10}\left(\frac{\quad}{\quad}\right) \; 0.4343(2.97\,Z^{1/2} R^{1/2} - 3.95\,Z\,T^{-1/2})$$

or $$\log_{10} \lambda = \log_{10}\left(\frac{v}{2R}\right) + 1.29\,Z^{1/2} R^{1/2} - 1.72\,Z\,T^{-1/2} \quad ...(8)$$

The changes in atomic number and nuclear radius are negligible when compared to the changes in energy. The first term is almost same for heavier nuclei. So Eq. (8) reduces to

$$\log_{10} \lambda = c + d\,T^{-1/2} \quad ...(9)$$

where "c" and "d" are constants.

Eq. (9) shows that the emitters having lesser decay constants emit α-particles of greater energy (T) which is the Geiger and Nuttal law.

Experimental verification of theory of α-decay

For a number of alpha-active nuclei, a plot of $\log_{10}\lambda$ *vs* $ZT^{-1/2}$ is shown in Fig. 31.14. The straight line has slope –1.72 as predicted by the theory. The intercept on the Y-axis gives the value of $\log_{10}\left(\frac{v}{2R}\right) + 1.29\,Z^{1/2} R^{1/2}$.

This can be used to determine the value of nuclear radius R. The nuclear radius calculated in this way comes out to be of the same order as obtained from scattering experiments.

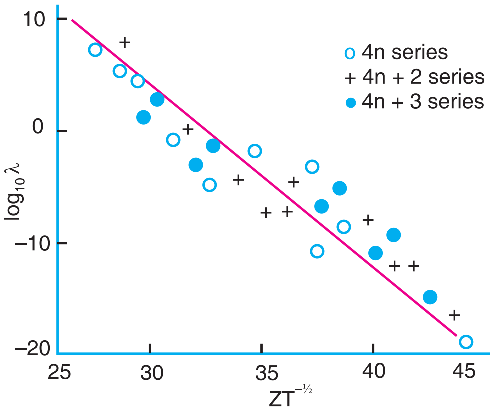


Fig. 31.14

Further this theory makes understandable the enormous variation in half-life with disintegration energy. Nuclei emitting most energetic α-particles are short lived and those emitting least energetic α-particles are long lived. This correlation between the half-life time (or disintegration constant λ) and the energy of the α-particle is predicted by Eq. (8).

BETA RAYS

31.16 The Nature of Beta Particles

The experiments of Becquerel and Curie showed that β-particles carried a negative charge. Kaufmann's experiments established that the charge to mass ratio (e/m) of these particles is the same as that for electrons. Hence, β-particles are identical with electrons.

31.17 Determination of e/m of Beta Particles : Kaufmann's Experiment

The apparatus used by Kaufmann is shown in Fig. 31.15. In a highly evacuated vessel, a speck of radium bromide (S) is placed. The charged particles from the source S are limited by a narrow slit L in a lead plate and are recorded on a photographic plate P. They are acted upon by a uniform electric field X parallel to the slit in the course of their travel from S to L. The electric field is produced by applying a high P.D. between two metal plates E_1 and E_2. The whole arrangement is placed between the pole-pieces of an electromagnet, so that the particles are acted upon by a uniform magnetic field B parallel to X. Hence the magnetic field produces a deflection in a direction perpendicular to that due to electric field. The comparatively heavy α-particles from the source produce a central spot (O) on P, since they are almost undeviated by the fields.

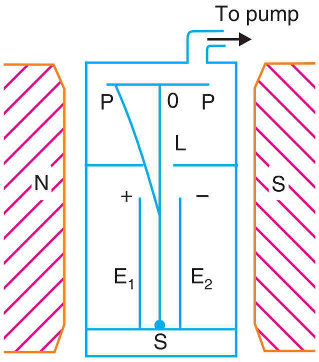


Fig. 31.15

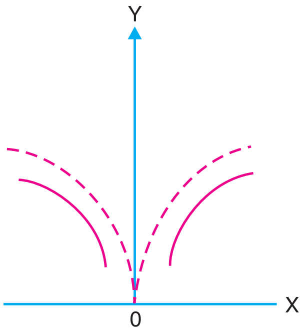


Fig. 31.16

The deflection due to the electric field $= x = \dfrac{XelL}{mv^2}$

where X = intensity of the electric field,

l = the length of the electrodes E_1 or E_2 over which the field acts

and L = the distance of the photographic plate from the field.

Due to the magnetic field, the particles travel in arcs of radius r given by

$$Bev = mv^2/r.$$

The deflection y due to the magnetic field is proportional to Be/mv and can be calculated from the geometry of the apparatus.

The source emits β-particles with a range of values of velocity v. Hence, the particles fall on the photographic plate along an arm of a parabola, each point on it corresponding to particles of a certain velocity (as in the case of Thomson's parabola method of positive ray analysis). When the electric or magnetic field is reversed, the other arm of the parabola is obtained. The parabolic traces obtained are shown in Fig. 31.16 by the continuous lines. The values of x and y are found for different points on the curve and hence e/m is calculated.

The traces actually obtained are found not to coincide with the parabolic curves that would be expected if e/m is the same for all velocities (shown by dotted lines in the figure). This shows that e/m is not constant but varies. For slow particles, e/m is found to be equal to that of an electron. The value of e/m decreases, as the velocity increases. This can be explained by the relativistic law of variation of mass with velocity $m = m_0 / \sqrt{1 - v^2/c^2}$. With the experimental value of e/m for the different values of v, Kaufmann calculated the value of e/m_0. The value of e/m_0 was found to be a constant for all values of v.

31.18 Bucherer's Experiment-Increase of Beta Particle Mass with Velocity

The source of β-particles is a small grain of radium fluoride R (Fig. 31.17). It is placed at the centre of two parallel discs A and B, 8 cm in diameter and 0.25 mm apart. A constant electric field X is maintained between the plates by establishing a high P.D. between them. The discs are surrounded by a coaxial cylindrical photographic film FF. The whole arrangement is enclosed in an evacuated chamber V. The whole apparatus is placed in a uniform magnetic field B acting parallel to the plane of the discs. A β-particle will only succeed in escaping from between the plates A, B when the magnetic and electric forces acting upon it are equal and opposite, since under any other field conditions it will be deflected to one or other of the plates. Considering a β-particle with velocity v moving along a radius making an angle θ with B, the condition for it to escape is

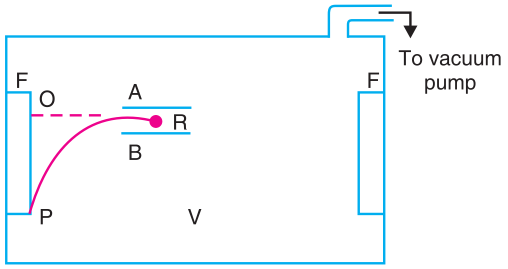


Fig. 31.17

$$Xe = Bev \sin\theta$$

or
$$v = \frac{X}{B\sin\theta} \qquad \text{...(1)}$$

Bucherer made $X/B = c/2$ in this experiment, so that $\sin\theta = c/2v$. According to the theory of relativity, v cannot exceed c. Hence θ cannot be less than 30° or greater than 150°. β-particles emerge from the space between the two plates only within these limiting directions. Along any direction θ between these limiting values, only those particles which have a unique value of v given by Eqn. (1) can emerge. β-particles with speeds other than this value are lost to the plates.

After emerging from the space between the plates, the β-particles are subjected to the magnetic field alone. The magnetic field acts on the component $v \sin\theta$ of the velocity and bends the particles into circular paths. The particles strike the film at the point P.

The deflection OP caused by the magnetic field is calculated as follows :

$$\left.\begin{array}{l}\text{The force due to the magnetic}\\ \text{field on the } \beta\text{-particles}\end{array}\right\} = Bev \sin\theta$$

$$\left.\begin{array}{l}\text{The acceleration acquired by}\\ \text{the } \beta\text{-particles in the field}\end{array}\right\} = \theta = \frac{Bev\sin\theta}{m}$$

The particles are subjected to the magnetic field for a time $t = a/v$

where a = distance from the edge of the discs to the film.

∴ The magnetic deflection
$$OP = \frac{1}{2}\alpha t^2$$
$$= \frac{1}{2}\frac{Bev\sin\theta}{m}\frac{a^2}{v^2} = \frac{1}{2}a^2\frac{Be\sin\theta}{mv}$$

$$= \left(\frac{e}{m}\right)\frac{a^2 B \sin\theta}{2v}$$

$$\therefore \qquad \frac{e}{m} = \frac{OP \times 2v}{a^2 B \sin\theta} \qquad \text{...(2)}$$

For any given value of θ, v is obtained from Eq. (1). OP and a can be readily measured. Hence e/m can be calculated for the particles of the velocity v.

By reversing X and B, a symmetrical pattern is obtained upon the photograhic film, which when opened out has the appearance shown in Fig. 31.18. From each of the points like P and P_1, the value of e/m can be calculated for a given value of v. In this way Bucherer determined accurately the values of e/m for several different values of v. Knowing the values of e/m and v, Bucherer calculated the value of e/m_0 from the relation

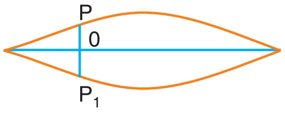


Fig. 31.18

$$\frac{e}{m} = \frac{e}{m_0}(1-\frac{v^2}{c^2})^{1/2}.$$

He found e/m_0 to be a constant, agreeing with the value for an electron at rest.

31.19 Beta Ray Spectra

The energies of β-particles from radio-active elements are determined by measuring the radii of curvature of their paths in a magnetic field of known flux density B. The circular path traversed by the β-particles of velocity v is governed by the relation

$$Bev = \frac{mv^2}{r}.$$

$$\therefore \qquad v = Br\,(e/m).$$

From the geometry of the arrangement, the radius of the circular path r can be found.

The value of e/m can be assumed. Hence, the velocity v can be calculated.

For particles moving with very high velocities, the kinetic energy of the particle,

$$E_k = mc^2 - m_0c^2 = m_0c^2\left[\frac{1}{(1-v^2/c^2)^{1/2}} - 1\right]$$

Example. *Calculate the velocity and the mass of a β-particle with a K.E. of 0.5 MeV. If a beam of such particles enters a uniform magnetic field, find the flux density required to cause them to traverse a circular arc of radius 0.1 m.*

Sol. Here,

$$E_k = 0.5\,MeV = 8.01 \times 10^{-14}\,J\,;\; m_0 = 9.11 \times 10^{-31}\text{ kg}$$

$$c = 3 \times 10^8\,ms^{-1},\; v = ?$$

$$8.01 \times 10^{-14} = (9.11 \times 10^{-31})(3 \times 10^8)^2\left[\frac{1}{[1-v^2/(3\times10^8)^2]^{1/2}} - 1\right]$$

$$\therefore \qquad v = 2.59 \times 10^8\,ms^{-1} = 0.863c.$$

$$\text{Mass of the } \beta\text{-particle} = m = \frac{m_0}{(1-v^2/c^2)^{1/2}} = \frac{9.11\times10^{-31}}{(1-0.863^2)^{1/2}} = 1.8\times10^{-30}\text{ kg}.$$

i.e., the mass of this β-particle is 1.98 m_0.

To find the magnetic flux density required, $Bev = mv^2/r$.

$$\therefore \qquad B = \frac{mv}{er} = \frac{(1.8 \times 10^{-30})(2.59 \times 10^{8})}{(1.6 \times 10^{-19})(0.1)} = 0.02914 \, weber/m^2.$$

31.20 Magnetic Spectrograph

The source is the radioactive substance *S* under study, coated on a fine wire. *A* and *B* are diaphragms used for limiting the beam of β-particles. *G* is a Geiger counter (Fig. 31.19). β-rays from the source at *S* are bent around by the magnetic field and focused on the aperture *O*. The number of β-particles entering the aperture *O* is counted at a given value of the magnetic field of intensity *B*. The intensity of the magnetic field is then changed to a new value and the number of β-particles entering the aperture *O* is again counted. In this manner, the velocity distribution of the β-particles is determined.

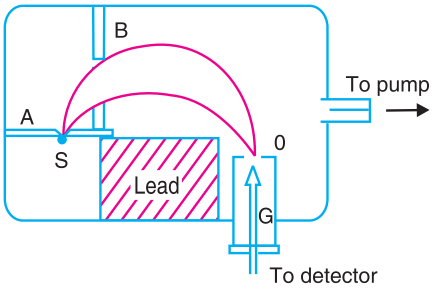
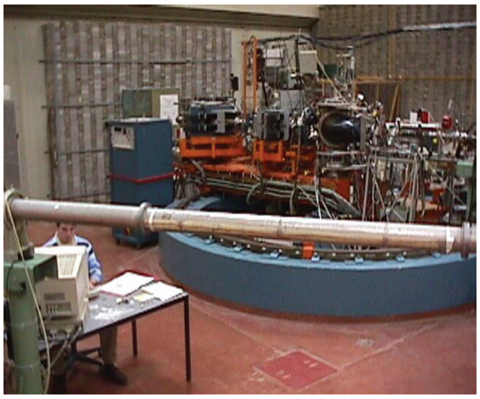


Fig. 31.19



This Setup is Designed for MeV ion Beam Analysis of thin Surface Layers with Ultra High Energy Resolution.

The continuous β-ray spectrum of radium *E* is shown in Fig. 31.20. The curve first rises to a maximum and then decreases to zero at a well defined velocity, which corresponds to the maximum velocity of the β-particles. Below the maximum energy (E_{max}), the β-ray particles give a continuous spectrum.

Results. The results of these measurements show that there are apparently two distinct types of β-ray spectrum, one a *sharp line spectrum* and the other a *continuous spectrum*. The lines correspond to β-particles emitted with discrete energies. The continuous spectrum is due to a continuous spread of energy among the emitted β-particles having a fixed maximum value of the energy. The upper energy limit of the continuous β-particle spectrum is characteristic of the radioactive atom concerned.

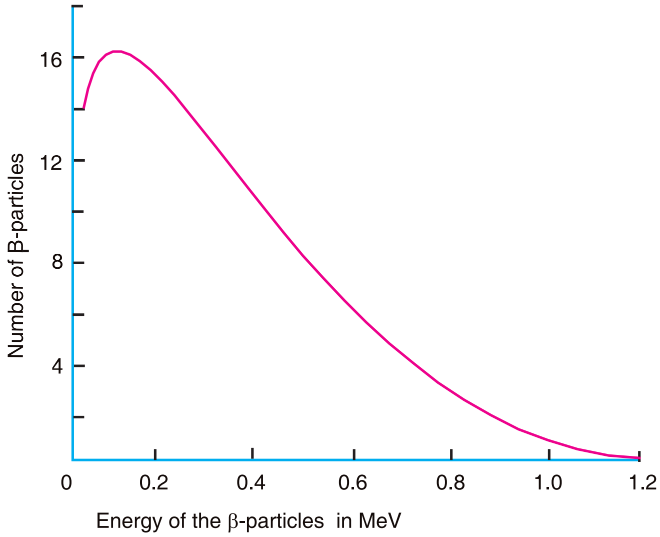


Fig. 31.20

31.21 Origin of the Line and Continuous Spectrum

It has definitely been shown that the sharp line spectra are due to electrons that have been ejected from the *K*, *L*, *M* and *N* shells of the atom by the process of *internal conversion*. The continuous β-ray spectrum is that produced by the electrons that have been ejected from the nuclei of radioactive atoms.

The β-ray spectrum of an element differs from the other spectra characteristic of the same element like optical, X-ray, α-ray and γ-ray. The β-ray spectrum is a continuous spectrum while the others are line spectra.

The explanation of the existence of the continuous energy spectrum for the β-particles presents some difficulties.

(1) When a nucleus emits a β-particle, a neutron in the nucleus changes to a proton. Hence the atomic number increases by unity and the mass number remains the same. Let M_1 = mass of the *neutral parent atom* of atomic number Z, M_2 = mass of *neutral daughter atom of atomic number* $(Z + 1)$, m = mass of the β-particle and e = Charge on the β-particle. Then, according to the principle of mass energy,

$$\left.\begin{array}{l}\text{Rest mass of}\\ \text{the parent nucleus}\end{array}\right\} = \text{Rest mass of daughter nucleus} + \text{Rest mass of electron} + \text{Energy of the electron.}$$

$$\text{or} \qquad (M_1 - Zm)\,c^2 = [M_2 - (Z + 1)\,m]\,c^2 + mc^2 + Q$$

$$\therefore \qquad \text{Energy of the electron} = Q = (M_1 - M_2)\,c^2.$$

Hence all the β-particles from a given radioactive substance must be emitted with the same K.E. But actual measurements show that only a few β-particles are emitted with this maximum value of energy. The majority of β-particles are emitted with smaller energies. What happens to the remaining energy?

(2) Another difficulty comes in the conservation of angular momentum. Every nucleus has an angular momentum (nuclear spin) which is an odd multiple of $\frac{1}{2}h/2\pi$ for nuclei of odd mass number and an even multiple of $\frac{1}{2}h/2\pi$ for nuclei of even mass number. The electron has an angular momentum $\frac{1}{2}h/2\pi$. In β-decay mass number remains unchanged. How is it possible for a nucleus of even mass number and therefore an integral spin to give rise to a daughter nucleus of the same mass number and also an integral spin and yet emit an electron of spin $\frac{1}{2}\,\frac{h}{2\pi}$? The same is the difficulty for a nucleus of odd mass number.

(3) There is also an apparent failure to conserve linear momentum in β-decay.

In 1930 Pauli proposed that if an uncharged particle of zero mass and spin $\frac{1}{2}$ is emitted in β-decay together with the electron, the energy, angular momentum and linear momentum discrepancies discussed above would be removed. The particle was named *neutrino.* It was supposed that neutrino carries off an energy equal to the difference between Q and the actual electron K.E. Subsequently it was found that there are two kinds of neutrino involved in β-decay, the neutrino itself (symbol ν) and the anti-neutrino (symbol $\overline{\nu}$). The reason neutrinos were not experimentally detected until recently is that their interaction with matter is extremely feeble. Lacking charge and mass, and not electromagnetic in nature, the neutrino can pass unimpeded through vast amounts of matter. A neutrino would have to pass through over 100 light-years of solid iron on the average before interacting.

31.22 The Neutrino Theory of Beta Decay

In 1934, Fermi developed a theory to explain the continuous β-ray spectrum. This theory is called neutrino theory of β-decay. According to this theory, a β-particle and a neutrino are created in the nucleus and both are emitted simultaneously. The *total energy* of these two particles is a constant which is equal to the end-point energy observed in the β-ray spectrum. This maximum energy is shared by the β-particle, the neutrino and also by the recoiling nucleus. The electron will carry the maximum energy when the energy of the neutrino is zero. In all other cases electron will carry an energy less than the maximum. The sum of the energies carried by the electron and the neutrino will always be the same. This energy may be shared by the two particles in any proportion. Hence it explains the continuous β-ray spectrum.

When the nucleon shifts from the neutron quantum state to the proton quantum state, electron and antineutrino are emitted. This process is represented by

$$n \rightarrow p + e^- + \bar{\nu}\,.$$

In ordinary beta decay it is an antineutrino that is emitted.

Positron emission corresponds to the conversion of a nuclear proton into a neutron, a positron, and a neutrino.

$$p \rightarrow n + e^+ + \nu.$$

Positron emission leads to a daughter nucleus of lower atomic number Z while leaving the mass number A unchanged. Thus negative and positive beta decays may be represented as

$$Z^{X^A} \rightarrow Z+1^{X^A} + e^- + \bar{\nu}$$

$$Z^{X^A} \rightarrow Z-1^{X^A} + e^+ + \nu$$

The electron, neutrino and product nucleus share among them the energy, angular momentum and linear momentum available from the nuclear transitions. Thus the neutrino theory of β-decay successfully explains the continuous energy spectrum of β-rays.

EXAMPLE : *Which of the following isobars would you expect to be* β^- *active and how would it decay ? Why ?* ${}_{28}Ni^{64} = 63.9280u$, ${}_{29}Cu^{64} = 63.928u$.

SOL. We know that for a β^- decay to be possible

$${}_ZM^A > {}_{Z+1}M^A + {}_{-1}e^0 \text{ (Nuclear masses).}$$

By adding Z electron masses to both sides, the nuclear masses are changed into isotopic masses. The same condition now appears as

$${}_ZM^A > {}_{Z+1}M^A \text{ (Atomic masses).}$$

The given isobars are β^- emitters, *i.e.*,

$${}_{28}Ni^{64} \rightarrow {}_{29}Cu^{64} + \beta^- \text{ and } {}_{29}Cu^{64} \rightarrow {}_{30}Zn^{64} + \beta^-$$

if ${}_ZM^A > {}_{Z+1}M^A$. Now mass of Cu^{64} is greater than that of Ni^{64}. Hence Ni^{64} is not a β^- emitter. Mass of Cu^{64} is greater than that of ${}_{30}Zn^{64}$ (63.92915u). Hence Cu^{64} is a β^- emitter.

K-electron capture. In an excited nucleus, there is a possibility of a proton turning into a neutron by absorbing an electron, instead of emitting a positron. In this case, an orbital electron of the parent nuclide, usually an electron in the K-shell which is closest to the nucleus, interacts with a proton in the nucleus, converting it into a neutron *i.e.*, the K-electron falls into the nucleus. The general basic equation is

$${}_ZX^A + {}_{-1}e^0 \rightarrow {}_{z-1}X^A + \nu.$$

The daughter nuclide is seen to be the same as that which would have been produced if positrons had been emitted. There is no question of any electron coming out from this process. The vacancy in the *K*-shell of the daughter nuclide resulting from *K*-capture is filled by the rearrangement of the *Z*–1 electrons in it. This leads to the emission of the characteristic X-rays of the daughter element. We also note that in β activity and in *K*-capture processes, the parent and the daughter nuclides form a pair of isobars. Some examples of the lighter radioactive nuclides which undergo electron capture are : Be^7, Al^{24}, Al^{26}, Ar^{37}, K^{40}, Ca^{41}, V^{48}, V^{49}, Cr^{51}, Fe^{55} and several others. In many cases, γ-radiation is also emitted. γ-radiation is emitted after electron capture if the nucleus is left in an excited state. In a few cases, both positron emission and electron capture occur. Example : Zn^{63}

$${}_{30}Zn^{63} \rightarrow {}_{29}Cu^{63} + {}_{1}e^{0}$$

$${}_{30}Zn^{63} + {}_{-1}e^{0} \rightarrow {}_{29}Cu^{63}$$

The three modes of beta radioactivity can be shown by the following reactions :

$${}_{0}n^{1} \rightarrow + {}_{1}p^{1} + {}_{-1}e^{0} \qquad (\textit{Negatron emission})$$

$${}_{+1}p^{1} \rightarrow {}_{0}n^{1} + {}_{+1}e^{0} \qquad (\textit{Positron emission})$$

$${}_{+1}p^{1} + {}_{-1}e^{0} \rightarrow {}_{0}n^{1} \qquad (\textit{Orbital electron capture})$$

GAMMA RAY SPECTRA

31.23 Introduction

γ-rays from radionuclides vary in wavelength from about 0.004 to 0.4 Å. γ-rays are of the same nature as X-rays and they accompany the emission of α and β-particles in many cases. The energy of the γ-rays is very large of the order of a few *MeV*. The γ-rays must be of nuclear origin. γ-rays may be regarded as very short X-rays and the methods of X-ray spectroscopy can be applied, with suitable modifications, for the determination of their wavelengths. The usual Bragg method may be adopted when the wavelength of the γ-rays exceeds 0.1 Å.

31.24 Determination of the Wavelength of Gamma Rays—

Du Mond Curved Crystal Spectrometer

The apparatus is shown in Fig. 31.21. The quartz diffracting crystal *C* is bent and clamped in a steel frame so that the diffracting planes meet, when extended, in a line *B* normal to the plane of the figure. The radius of curvature of the crystal is then equal to the diameter of the focusing circle *F*. The γ-ray source is kept on the focusing circle at point *R*. A detector like a scintillation counter is placed behind the crystal.

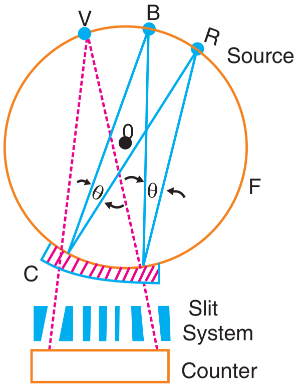


Fig. 31.21

The counting rate is measured as a function of the source position (angle θ). *R* is moved on a circle and the output of the detector is plotted against θ. At the diffraction maxima, peaks are found. From the value of θ corresponding to the peaks, the value of λ is calculated from Bragg's formula. The advantage of this method is that it avoids the necessity of collimating the incident γ-rays and makes full use of the whole of the aperture of the crystal.

Disadvantages (1) The measurements become more difficult and less precise as the energy of the γ-rays increases and the wavelength decreases.

(2) This method requires highly active sources which are often not available.

31.25 Origin of γ-rays

The large values of energy associated with γ-rays show that they must be of nuclear origin. γ-ray spectrum consists of sharp lines and this indicates the existence of a number of energy levels in the nucleus. Stable nuclides are usually in the state of least energy or ground state, but they can be excited by particle or photon bombardment. Hence nuclei can exist in states of definite energies, just as atoms can. An excited nucleus is denoted by an asterisk (*) after its usual symbol. Thus ${}_{38}\mathrm{Sr}^{87*}$ refers to ${}_{38}\mathrm{Sr}^{87}$ in an excited state. One way an excited nucleus can return to the ground state is by the emission of γ-rays. γ-ray decay is represented schematically by

$$({}_Z X^A)^* \rightarrow {}_Z X^A + \gamma .$$

The star (*) indicates an excited nucleus, and both the daughter and the parent have the same structure of nuclear particles. If E^* is the energy associated with the excited state and E is the energy of the ground state, then the γ-rays have an energy

$$h\nu = E^* - E$$

where ν is the frequency of the emitted γ-ray.

A simple example of the relationship between energy levels and decay schemes is shown in Fig. 31.22 which pictures the β-decay of ${}_{12}\mathrm{Mg}^{27}$ to ${}_{13}\mathrm{Al}^{27}$. The half-life of the decay is 9.5 minutes, and it may take place to either of the two excited states of ${}_{13}Al^{27}$. The resulting ${}_{13}Al^{27*}$ nucleus then undergoes one or two gamma decays to reach the ground state.

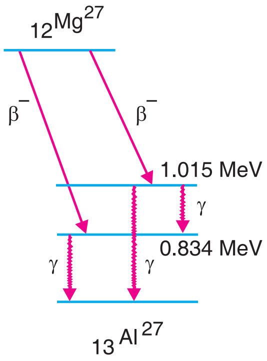


Fig. 31.22

Most excited nuclei have very short half-lives against γ-decay, but a few remain excited for as long as several hours. A long lived excited nucleus is called an *isomer* of the same nucleus in its ground state. The excited nucleus ${}_{38}\mathrm{Sr}^{87*}$ has a half-life of 2.8 hours and is an isomer of ${}_{38}\mathrm{Sr}^{87}$.

31.26 Nuclear Isomerism

There are nuclei which have the same atomic and mass numbers (same Z and same A) but differ from one another in their nuclear energy states and exhibit differences in their internal structure. These are called *nuclear isomers.*

The existence of nuclear isomers is called *nuclear isomerism.* The excited nucleus ${}_{38}\mathrm{Sr}^{87*}$ is an isomer of ${}_{38}\mathrm{Sr}^{87}$. The difference between the nuclear isomers is attributed to a difference of nuclear energy states. One isomer represents the nucleus in its ground state, whereas the other is the same nucleus in an excited state of higher energy. The phenomenon of nuclear isomerism was discovered by *O. Hahn* in 1921. He found that UX_2 and UZ both have the same atomic number and the same mass number but have different half-lives and emit different radiations. UX_2 has 0.394 MeV more energy in its nucleus than UZ. Both these nuclei are formed out of UX_1 by β–decay. UX_2 has half-life of 1.17 minutes and UZ has a half-life of 6.7 hours. The higher energy isomer UX_2 may directly decay to UII by β emission with a half-life of 1.17 minutes, or it may first come to the lower energy isomer by emitting a γ-ray of energy 0.394 MeV and then decay to UII by β-emission with a half-life of 6.7 hours.

Nuclear isomerism has also been detected in artificial radioactive substances. Many *isomeric pairs* have been produced by bombarding radionuclides with neutrons.

31.27 Internal Conversion

When a nucleus passes from a higher excited state to the ground state, the difference in energy of the two states is emitted as a γ-ray. As an alternative to γ-decay, an excited nucleus, in some cases, may return to its ground state by giving up its excitation energy to one of the orbital electrons around it. The emitted electron has a K.E. equal to the lost nuclear excitation energy minus the binding energy of the electron in the atom. *i.e.*,

K.E. of the ejected electron $= E_e - W$.

where E_e = the available excitation energy and

W = binding energy of the ejected electron in its shell of origin.

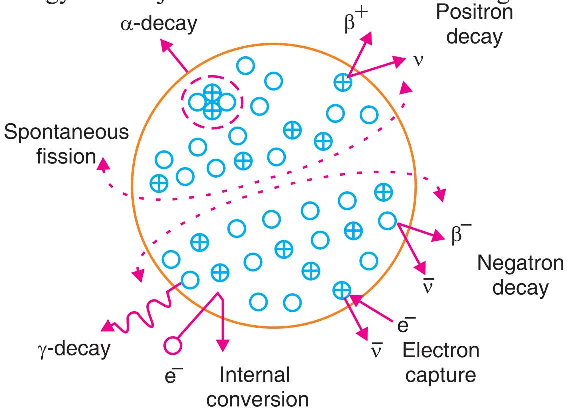


Fig. 31.23

This process is called *internal conversion.* The emitted electron is called a *conversion electron.* Thus internal conversion and emission of a γ-ray from the nucleus are two alternate ways of accomplishing the same nuclear transition. The internal conversion is not a *two step* process in which a γ-ray photon is first emitted and then it knocks out an orbital electron. It is in better accord with experiment to regard internal conversion as representing a direct trnasfer of excitation energy from a nucleus to an orbital electron. Hence, internal conversion is a *single step* process in which the excited nucleus interacts ejected electron (β-particle) has discrete values. Therefore, the corresponding β-particle energy spectrum is a line spectrum having discrete energies.

Fig. 31.23 illustrates the various kinds of disintegration processes that radioactive nuclei may ungergo. The nucleus is represented as an assembly of protons and neutrons. A proton is indicated by a cross, and a neutron by an open circle.

31.28 Mossbauer Effect

The Mossbauer Effect. The emission of gamma rays is generally accompanied by the emission of an α or β-particle. If after the emission of an α or β-particle the product nucleus is left in an excited state, it reaches the ground state by releasing or emitting photons called γ-rays. When a nucleus emits a photon it recoils in the opposite direction. This reduces the energy of the γ-rays from its usual transition energy E_0 to $E_0 - R$, where R is the recoil energy.

In certain crystals, the atoms are tightly bound to the lattice. When such an atom emits a γ-ray, the entire crystal and not merely the emitting atom recoils as a whole. In this case the recoil energy is not transferred to *one* emitting atom, but to a *group* of atoms. The whole mass of crystal — not just the recoiling particle — takes part in the recoil. This implies that the recoiling mass is extremely large and the *recoil velocity is almost zero.* Now the γ-ray photon has the transition energy E_0. The emission of γ-rays without recoil, and therefore without reduction of the γ-ray energy, is called the *Mössbauer effect.*

Experimental arrangement. Fig. 31.24 (*a*) shows the experimental arrangement.

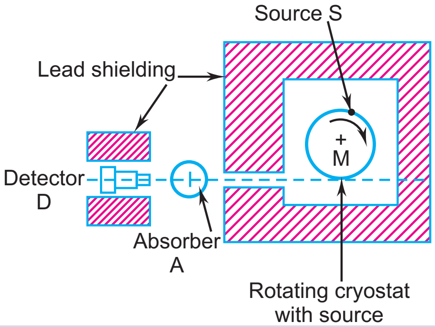


Fig. 31.24 (a)

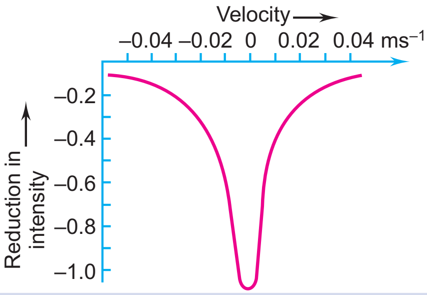


Fig. 31.24 (b)

- *S* is the source at the rim of rotating cryostat. The source Os^{191} decays by β^- emission to an excited state of stable Ir^{191} with energy 0.129 MeV. *M* is the region in which the source is seen from the scintillation detector.
- *A* is the absorber in a cryostat. The absorber is Ir^{191} (a foil of natural crystalline irridium).
- *D* is a scintillation detector. Lead blocks are used for shielding purposes.

Procedure. The recoil-free γ-rays from the source *S* are collimated through a channel in a *Pb*-block that encloses the source. The radiation falls on the absorber *A*. Both the source and the absorber are placed in two cryostats maintained at a temperature of 88 K. The cooling reduces the thermal broadening to an enormous extent. The γ-photons transmitted by *A* are detected with the scintillation counter *D*.

The cryostat containing *S* is mounted on a lathe carriage and is set into rotation. The detector is turned on only when the source is in line with the absorber *A*. The Doppler shift in the frequency of γ-photon offsets the resonance absorption by the stationary absorber *A*.

Resonance Curve. Fig. 31.24 (*b*) shows the *variation of γ-intensity with relative velocity of source and absorber.* The faster the motion of the source, greater is the mismatch between the emission and absorption frequencies and so less is the absorption in the foil or more is the number of γ-photons reaching the counter *C*. The resonance absorption of γ-rays is so sharp that even a relative speed of 0.02 ms^{-1} reduces the absorption by more than 50% at 88 K. The curve in Fig. 31.24(*b*) corresponds to the theoretical resonant nuclear absorption for 129 kV γ-photons from ^{191}Ir, based on a natural width of 5×10^{-6} eV for its excited state. The experimental points of Mössbauer fall close to this curve.

Review Question

What is Mössbauer effect? Describe experimental set up to study this effect.

(*Mumbai University, October 2010*).

FUNDAMENTAL LAWS OF RADIOACTIVITY

31.29 Soddy Fajan's Displacement Law

(1) *In all known radioactive transformations either an α or a β-particle* (*i.e.*, *never both or more than one of each kind*) *is emitted by the atom.*

(2) *When a radioactive atom emits an α particle, a new atom is formed whose mass number is less by four units and atomic number less by two units than those of the parent atom.*

$${}_{92}U^{238} \xrightarrow{-\alpha} {}_{90}Th^{234}$$

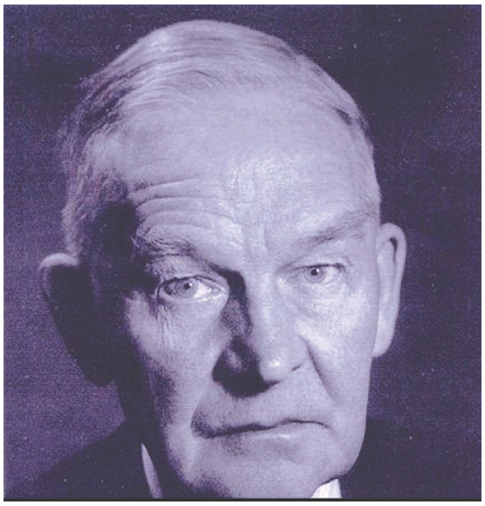
(3) *When a radioactive atom emits a β-particle, the new atom formed has the same mass number but the atomic number is increased by one unit.*

$${}_{90}Th^{234} \xrightarrow{-1^{e0}} {}_{91}pa^{234}$$

Natural Radioactive Series

With the help of the displacement law, one can determine easily the mass and atomic numbers of the different elements in the successive radioactive changes, if the mass and atomic numbers of the parent are known. The different atoms produced at the different stages of disintegration will naturally form a *series.*

The first member is called the *parent*, the intermediate members are called *daughters* and the final stable member is called the *end product*. (1) Uranium series (2) Actinium series (3) Thorium series and (4) Neptunium series are four such series known at present. Each of the series consists of a succession of daughter elements which are all derived from a single parent nuclide. An explanation can be given as to why there are exactly four series. We know that alpha decay reduces the mass number of a nucleus by 4. Thus the nuclides, whose mass numbers are all given by $A = 4n$, where n is an integer, can decay into one another in descending order of mass number. All radioactive nuclides whose mass numbers obey the equation $A = 4n$ are said to be members of the $4n$ series. Radioactive nuclides whose mass numbrs obey the equation $A = 4n + 1$ belong to the $(4n + 1)$ series. Similarly radioactive nuclides whose mass numbers satisfy the equation $A = 4n + 2$ and $A = 4n + 3$ belong to the $(4n + 2)$ series and $(4n + 3)$ series respectively.



Frederick Soddy.

TABLE 31.1 Four Radioactive Series

Mass numbers	*Series*	*Parent*	*Half-life, years*	*Stable end product*
$4n$	Thorium	${}_{90}Th^{232}$	1.39×10^{10}	${}_{82}Pb^{208}$
$4n + 1$	Neptunium	${}_{93}Np^{237}$	2.25×10^{6}	${}_{83}Bi^{209}$
$4n + 2$	Uranium	${}_{92}U^{238}$	4.51×10^{9}	${}_{82}Pb^{206}$
$4n + 3$	Actinium	${}_{92}U^{235}$	7.07×10^{8}	${}_{82}Pb^{207}$

Table 31.1 is a list of the names of four important radioactive series, their parent nuclides and the half-lives of these parents, and the stable daughters which are end products of the series.

31.30 Law of Radioactive Disintegration

Let N be the number of atoms present in a particular radioelement at a given instant t. Then the rate of decrease $-dN/dt$ is proportional to N.

$$\therefore \qquad -\frac{dN}{dt} = \lambda N \qquad \text{...(1)}$$

Here λ is a constant known as the *disintegration constant* or *decay constant* of the radioactive element. *It is defined as the ratio of the amount of the substance which disintegrates in a unit time to the amount of substance present.*

Eq. (1) can be written as $\frac{dN}{N} = -\lambda\, dt.$

Integrating, $\log_e N = -\lambda\, t + C$ ...(2)

Let the number of radioactive atoms initially present be N_0.

Then, when $t = 0$, $N = N_0$. $\therefore$ $\log_e N_0 = C$.

Substituting for C in (2), we get, $\log N = -\lambda\, t + \log N_0$

or
$$\log_e \frac{N}{N_0} = -\lambda t$$

or
$$N = N_0 e^{-\lambda t} \quad ...(3)$$

This equation shows that the number of atoms of a given radioactives substance decreases exponentially with time (Fig. 31.25).

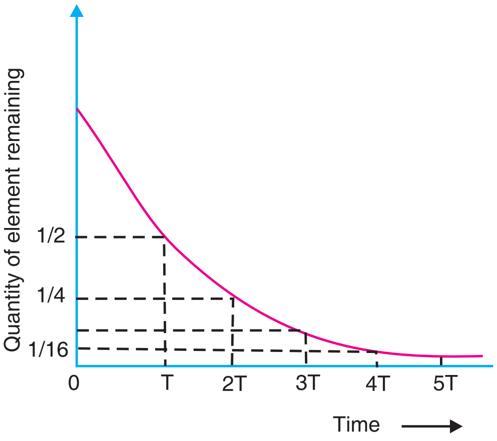


Fig. 31.25

Definition of half-life period. *The half-life period of a radioactive substance is defined as the time required for one-half of the radioactive substance to disintegrate.*

Value of half-life period

We know the relation $N = N_0\, e^{-\lambda t}$.

If $T_{1/2}$ be the half-life period, then at

$$t = T_{1/2},\ N = N_0/2.$$

$$\therefore \quad \frac{N_0}{2} = \log_e 2 \text{ or } e^{\lambda T_{1/2}} = 2$$

$$\text{or } \lambda T_{1/2} = \log_e 2 \text{ or } T_{1/2} = \frac{\log_e 2}{\lambda}$$

$$\therefore \quad T_{1/2} = \frac{\log_e 2}{\lambda} = \frac{0.6931}{\lambda}.$$

Example 1. *The half-value period of radium is* 1590 *years. In how many years will one gram of pure element* (*a*) *lose one centigram, and* (*b*) be *reduced to one centigram* ?

Sol. Here, half-life period of radium = $T_{1/2}$ = 1590 years.

$$\therefore \quad \text{Radioactive constant} = \lambda = \frac{0.6931}{T_{\frac{1}{2}}} = \frac{0.6931}{1590}.$$

(*a*) Let t be the time in which one gram of radium loses one centigram (0.01g).

$\therefore$ Radium left behind = 1 – 0.01 = 0.99 gram.

Now, $N = N_0\, e^{-\lambda t}$ or $\log_e N = \log_e N_0 - \lambda\, t$

or
$$\lambda t = \log_e \left(\frac{N_0}{N}\right)$$

$$\therefore \quad t = \frac{1}{\lambda}\log_e\left(\frac{N_0}{N}\right) = \frac{1590}{0.6931}\log_e\left(\frac{1}{0.99}\right)$$

$$= 23.25 \text{ years.}$$

(*b*) Here, $N = 0.01$ gram; $N_0 = 1$ gram; $t = ?$

$$t = \frac{1}{\lambda}\log_e\left(\frac{N_0}{N}\right) = \frac{1590}{0.6931}\log_e\left(\frac{1}{0.01}\right) = 10560 \text{ years.}$$

Example 2. *1 gram of radium is reduced by* 2.1 *mg in* 5 *years by* α *-decay. Calculate the half-life period of radium.*

Sol. Initial mass of radium = 1 gram

Mass of radium left behind after 5 years = 1 – 0.0021 = 0.9979 g.

We have, $N = N_0\, e^{-\lambda t}$; Here, $\frac{N}{N_0} = \frac{0.9979}{1}$; $t = 5$ years.

$$\therefore \qquad 0.9979 = e^{-5\lambda} \text{ or } e^{5\lambda} = \frac{1}{0.9979}$$

$$\text{or} \qquad 5\lambda = \log_e\left(\frac{1}{0.9979}\right)$$

$$\therefore \qquad \lambda = \frac{1}{5}\log_e\left(\frac{1}{0.9979}\right) = \frac{2.3026}{5}\log_{10}\left(\frac{1}{0.9979}\right)$$

$$= 41.45 \times 10^{-5} \text{ per year.}$$

$$T_{1/2} = \frac{0.6931}{\lambda} = \frac{0.6931}{41.45\times10^{-5}} = 1672 \text{ years.}$$

Example 3. *Calculate the time required for* 10% *of a sample of thorium to disintegrate. Assume the half-life of thorium to be* 1.4×10^{10} *years.*

Sol. We have, $N = N_0\, e^{-\lambda t}$. Here $N = 0.9\, N_0$.

$$\lambda = \frac{0.6931}{T_{1/2}} = \frac{0.6931}{1.4\times10^{10}}\,a^{-1}.$$

$$\therefore \qquad 0.9\, N_0 = N_0\, e^{-\lambda t} \quad \text{or} \quad \lambda t = \log_e\left(\frac{1}{0.9}\right) \text{ or } t = \frac{1}{\lambda}\log_e\left(\frac{1}{0.9}\right)$$

$$t = \frac{1.4\times10^{10}}{0.6931}\times\log_e 1.111 = \frac{1.4\times10^{10}}{0.6931}\times 2.302\times\log_{10}1.111$$

$$= 2.1 \times 10^{9} \text{ years.}$$

31.31 The Mean life

It is not possible to predict which atom of a radioactive substance will disintegrate at any instant. The atom which disintegrates first has zero life and that disintegrates last has infinite life. Thus the life of every atom is different and the actual lives of the various atoms range from zero to infinity.

$$\therefore \quad \text{The mean-life of a radioactive element} = \frac{\text{Sum of the lives of all the atoms}}{\text{Total number of atoms}}.$$

Definition. *The mean-life of a radioelement is defined as the ratio of the total life time of all the radioactive atoms to the total number of such atoms in it.*

Value of mean-life. Let N_0 be the total number of radioactive atoms in the beginning. Let N be the number of atoms of that element after time t. Then $N = N_0\, e^{-\lambda t}$. Let dN be the number of atoms disintegrating between time t and $t + \delta t$. These dN atoms have had a life between t and $(t + \delta\, t)$. Since δt is very small, each of these atoms had a life of t.

$\therefore$ total life of dN atoms = $(dN)t$.

The possible life of any of the total number N_0 radioactive atoms varies from 0 to ∞.

$\therefore$ Total life time of all N_0 atoms = $\int_0^\infty t\,dN$

$$\text{Now, mean life} = \overline{\text{T}} = \frac{\text{total life} - \text{time}}{\text{total number of atoms}}$$

$$= \frac{\int_0^\infty t\,dN}{N_0}$$

Now, $\quad N = N_0 e^{-\lambda t}$

$\therefore \quad \frac{dN}{dt} = -\lambda N_0 e^{-\lambda t}$

or $\quad dN = -\lambda N_0 e^{-\lambda t}\,dt$

$\quad dN = \lambda N_0 e^{-\lambda t}\,dt$

(Leaving the negative sign which merely indicates the decrease in the number of atoms with time).

Hence, $$\overline{T} = \frac{\int_0^\infty t\lambda N_0 e^{-\lambda t}\,dt}{N_0} = \lambda\int_0^\infty t\,e^{-\lambda t}\,dt$$

Integrating by parts,

$$\overline{T} = \lambda\left[\frac{te^{-\lambda t}}{-\lambda} - \int\frac{e^{-\lambda t}\,dt}{-\lambda}\right]_0^\infty$$

$$= \lambda\left[\frac{te^{-\lambda t}}{-\lambda} - \frac{e^{-\lambda t}}{-\lambda^2}\right]_0^\infty = \lambda\left(\frac{1}{\lambda^2}\right) = \frac{1}{\lambda}$$

$\therefore \quad \overline{T} = \frac{1}{\lambda}.$

Thus the mean life (T) of a radioactive substance is the reciprocal of the decay constant (λ).

EXAMPLE 1. *1 gram of a radioactive substance disintegrates at the rate of* 3.7×10^{10} *disintegrations per second. The atomic weight of the substance is 226. Calculate its mean life.*

SOL. Number of atoms disintegrated in one second = 3.7×10^{10}

The mass of the substance disintegrated in one second = $\frac{(3.7 \times 10^{10}) \times 226}{6.02 \times 10^{26}}$

$= 1.389 \times 10^{-14}$ kg

Here, $N = 1$ g $= 10^{-3}$ kg; $\quad -dN/dt = 1.389 \times 10^{-14}$ kg

We have, $-dN/dt = \lambda N$

or $$\lambda = \frac{-dN/dt}{N} = \frac{1.389 \times 10^{-14}}{10^{-3}} = 1.389 \times 10^{-11}\,s^{-1}$$

∴ Mean life $\bar{T} = \frac{1}{\lambda} = \frac{1}{1.389 \times 10^{-11}}\ s = 2282\ years$.

Example 2. *The disintegration constant* λ *of a radioactive element is* 0.00231 *per day. Calculate its half-life and average life.*

Sol. Here, $\lambda = 0.00231$ per day.

(*i*) Half-life period $= T_{1/2} = \frac{0.693}{\lambda} = \frac{0.693}{0.00231} = 300\,\text{days}$

(*ii*) Average life period $= \bar{\text{T}} = \frac{1}{\lambda} = \frac{1}{0.00231} = 432.9\,\text{days}$.

31.32 Measurement of Decay Constants

The three constants λ, $T_{1/2}$ and $\bar{T}$ of a radioelement are interrelated. $T_{1/2} = \frac{0.693}{\lambda}$ and $\bar{\text{T}} = \frac{1}{\lambda}$.

The decay constant λ is determined by experiment. Different methods are employed according as λ is large or small. If λ is known, $T_{1/2}$ and $\bar{T}$ can be calculated.

(*i*) Elements of short period (large λ). The decay curve is drawn with experimentally measured values of intensity and time. The logarithm of intensity is plotted against time. A straight line graph is obtained representing the relation

$\log N = \log N_0 - \lambda t$ or $y = a - \lambda t$

where $y = \log N$ and $a = \log N_0$ (constant).

$\frac{dy}{dt} = -\lambda$. Thus the slope of the straight line gives λ.

(*ii*) Elements of long period (small λ). Each disintegrating atom emits only one α-particle. Hence the number of α-particles counted must be equal to the number of atoms disintegrating.

We have, $-\frac{dN}{dt} = \lambda N$ or $\lambda = \frac{-dN/dt}{N}$

i.e., λ is equal to the ratio of the atoms disintegrating per unit time to the total number of atoms present.

Example. *Geiger and Rutherford found by the method of scintillations that thorium emits* 4500 *α-particles per second per gram.* 1 *gram of thorium contains* 2.61×10^{21} *atoms. Calculate its half-life.*

Sol. $\lambda = \frac{-dN/dt}{N} = \frac{4500}{2.61 \times 10^{21}} = 1.72 \times 10^{-18}\ s^{-1} = 5.44 \times 10^{-11}\ \text{year}^{-1}$.

∴ $T_{1/2} = \frac{0.693}{\lambda} = \frac{0.693}{5.44 \times 10^{-11}} = 1.28 \times 10^{10}$ years.

31.33 Units of Radioactivity

Generally we express the activity of a radioactive substance in terms of *curie* and its submultiples *millicurie* (*mCi*) and micro curie (μ*Ci*). The curie is defined as the quantity of a radioactive substance which gives 3.70×10^{10} disintegrations / second. This is the number of disintegrations per second per gram of radium. Another *unit rutherford* (*Rd*) is also used. It is defined as the quantity of a radioactive substance which gives 10^6 disintegrations / second. It is to be noted that for a substance with very short half-life period, very little of the substance is required for 1 curie of activity. On the other hand, for the substance with very long half-life peroid, a very large quantity of the substance is required for 1 curie of activity.

Example. *Calculate the weight in kg of one curie of Ra B* (Pb^{214}) *from the half-life of* 26.8 *minutes.*

Sol. 1 curie = 3.7×10^{10} disintegrations / sec.

Here, $-dN/dt = \lambda N = 3.7 \times 10^{10}$

or
$$N = \frac{3.7 \times 10^{10}}{\lambda} = \frac{3.7 \times 10^{10} \times T_{1/2}}{0.693}$$

∴
$$N = \frac{(3.7 \times 10^{10})(26.8 \times 60)}{0.693} = 8.585 \times 10^{13} \textit{ atoms.}$$

From Avogadro's hypothesis, 6.023×10^{26} atoms are contained in 214 kg of *RaB*.

8.585×10^{13} atoms are contained in $\dfrac{214 \times 8.585 \times 10^{13}}{6.023 \times 10^{26}}$ kg.

$$= 3.1 \times 10^{-11} \text{ kg.}$$

Activity. The *activity* of a sample of any radioactive nuclide is the rate at which the nuclei of its constituent atoms decay. If N is the number of nuclei present in the sample at a certain time, its activity R is given by $R = -dN/dt$. The S.I. unit of activity is named after Henri Becquerel.

1 becquerel = 1 Bq = 1 event/s.

1 MBq = 10^6 Bq and 1 GBq = 10^9 Bq.

Example. *Find the activity of* 1 *mg* (10^{-6}*kg*) *of radon.*

Sol. The decay constant of randon is

$$\lambda = \frac{0.693}{3.8\,d} = \frac{0.693}{3.8 \times 86400\,s} = 2.1 \times 10^{-6}\, s^{-1}.$$

The No. of atoms in 10^{-6} kg of Rn^{222} $= N = \dfrac{(6.023 \times 10^{26}) \times 10^{-6}}{222} = 2.7 \times 10^{18}$

∴
$$\begin{aligned} R &= \lambda N = (2.1 \times 10^{-6}) \times (2.7 \times 10^{18}) \\ &= 5.7 \times 10^{12} \text{ events / s} \\ &= 5.7 \times 10^{3}\ GBq = 153\ Ci \end{aligned}$$

In this case the "events" are alpha decays.

31.34 Law of Successive Disintegration

Chain disintegration means a process in which a radioactive substance disintegrates to form a new substance, which disintegrates to form another new substance and so on. Consider a substance A which decays to form a substance B. Let the substance B decay to form a substance C and so on.

$$A \to B \to C \to \ldots X \text{ (stable).}$$

In a radioactive series, any two adjacent elements are considered as *parent* and *daughter*. Evidently in a series, the parent of the following element will be daughter of the preceding one.

At time $t = 0$, the number of initial atoms in $A = N_0$

and the number of initial atoms in $B = 0$

At time t, let the number of atoms in $A = N_1$

and the number of atoms in $B = N_2$.

Let λ_1 and λ_2 be the decay constants of A and B respectively. Every time an atom of A disappears, an atom of B is produced.

∴ Rate of formation of daughter $B = \lambda_1 N_1$.

The rate at which B decays $= \lambda_2 N_2$.

∴ The net increase in the number of B atoms

$$\frac{dN_2}{dt} = \lambda_1 N_1 - \lambda_2 N_2$$

Hence
$$\frac{dN_2}{dt} = \lambda_1 N_0 e^{-\lambda_1 t} - \lambda_2 N_2 \text{ (since } N_1 = N_0 e^{-\lambda_1 t}\text{)}.$$

∴
$$\frac{dN_2}{dt} + \lambda_2 N_2 = \lambda_1 N_0 e^{-\lambda_1 t}$$

Multiplying both sides by the integrating factor $e^{\lambda_2 t}$,

$$\frac{dN_2}{dt} e^{\lambda_2 t} + \lambda_2 N_2 e^{\lambda_2 t} = \lambda_1 N_0 e^{(\lambda_2 - \lambda_1)t}$$

or
$$\frac{d}{dt}[N_2 e^{\lambda_2 t}] = \lambda_1 N_0 e^{(\lambda_2 - \lambda_1)t}$$

Integrating,
$$N_2 e^{\lambda_2 t} = \frac{\lambda_1}{\lambda_2 - \lambda_1} N_0 e^{(\lambda_2 - \lambda_1)t} + C \quad \text{...(1)}$$

The constant of integration C can be determined from the initial conditions.

i.e., when $t = 0, N_2 = 0$; ∴ $C = \dfrac{-\lambda_1 N_0}{\lambda_2 - \lambda_1}$

Substituting the value of C in (1),

$$N_2 e^{\lambda_2 t} = \frac{\lambda_1 N_0}{\lambda_2 - \lambda_1}[e^{(\lambda_2 - \lambda_1)t} - 1]$$

or
$$N_2 = \frac{\lambda_1 N_0}{\lambda_2 - \lambda_1}[e^{-\lambda_1 t} - e^{-\lambda_2 t}] \quad \text{...(2)}$$

Radioactive equilibrium :

Case (*i*). Secular or permanent equilibrium. Suppose $T_1 >> T_2$ with $T_1 \approx \infty$ and $T_2 = 0$. (*i.e.*, the half-life of A is very much longer than that of B).

Then, $\lambda_1 << \lambda_2$ and $\lambda_1 \approx 0$. In equation (2), $e^{-\lambda_1 t} \to 1$

∴
$$N_2 = \frac{N_0 \lambda_1}{\lambda_2}(1 - e^{-\lambda_2 t}) \quad \text{...(3)}$$

After an appreciably long time, $e^{-\lambda_2 t}$ becomes negligible and

$$N_1 \approx N_0.$$

∴
$$N_1 \lambda_1 = N_2 \lambda_2 \quad \text{...(4)}$$

This relation shows that at equilibrium, the rate of decay of any radioactive product is just equal to its rate of production from the previous member of the chain. Now the daughter is said to be in *secular* or *permanent equilibrium* with the parent. An example of this is provided by the decay of radium into radon according to the equation.

$$Ra^{226} \rightarrow Rn^{222} + \alpha.$$

The half-life period of *Ra* (1590 years) is very large compared with that of *Rn* (3.8 days). Fig. 31.26 shows the experimental decay and recovery curves. After a time, long enough compared with its mean life, *Rn* is in permanent equilibrium with *Ra*. Its amount becomes constant as shown by the sum of the ordinates of the two curves which becomes constant beyond a certain value of time indicated by the point of intersection.

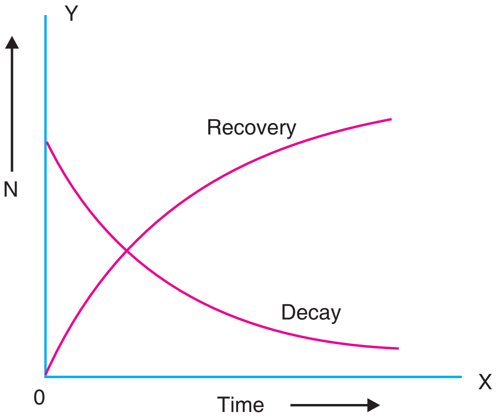


Fig. 31.26

Case (*ii*). Transient equilibrium. Suppose $\lambda_1 << \lambda_2$ but $\lambda_1 \neq 0$ *i.e.*, the decay constant λ_1 which is very small compared with λ_2, is not small enough to be ignored. In this case, after a sufficiently long time, $e^{-\lambda_2 t}$ becomes negligible compared with $e^{-\lambda_1 t}$. Hence

$$N_2 = \frac{N_0 \lambda_1}{\lambda_2 - \lambda_1} e^{-\lambda_1 t} = \frac{\lambda_1 N_1}{\lambda_2 - \lambda_1}$$

or $$\frac{N_2}{N_1} = \frac{\lambda_1}{\lambda_2 - \lambda_1} \qquad ...(5)$$

In this case, both *A* and *B* decay while the ratio N_2 / N_1 remains constant.

31.35 Radioactive Dating : The Age of the Earth

The age of the earth is estimated from the relative abundance of the two isotopes of uranium, U^{238} and U^{235}. The half-periods of U^{238} and U^{235} are 4.5×10^9 years and 7×10^8 years respectively. Assume that at the beginning when the earth was formed the proportions of the two isotopes were equal. The present relative abundance of U^{238} to U^{235} in natural uranium is 99.3% to 0.7%.

$$\therefore \quad \frac{N_1}{N_2} = \frac{99.3}{0.7} = \frac{N_0 e^{-\lambda_1 t}}{N_0 e^{-\lambda_2 t}} = e^{(\lambda_2 - \lambda_1)t}$$

where $$\lambda_1 = \frac{0.6931}{4.5 \times 10^9} \text{ and } \lambda_2 = \frac{0.6931}{7 \times 10^8}$$

$$\therefore \quad \log_e \left(\frac{99.3}{0.7} \right) = (\lambda_2 - \lambda_1)t$$

$$\therefore \quad t = \frac{1}{\lambda_2 - \lambda_1} \log_e \left[\frac{99.3}{0.7} \right]$$

$$= \frac{1}{\left(\frac{0.6931}{7 \times 10^8} \right) - \left(\frac{0.6931}{4.5 \times 10^9} \right)} \log_e \left(\frac{99.3}{0.7} \right)$$

$$= 5.93 \times 10^9 \text{ years.}$$

This value agrees nearly with that given by astronomical evidence for the age of the universe.

Radioactive Decay Model.

Dating by radioactive decay. The decay of radioactive elements is independent of the physical and chemical conditions imposed on them. Although the decay of an individual particle from a given nucleus is a random process, the gross decay of the many nuclei in a given sample provides a very convenient way of measuring times. In 1913 Joly and Rutherford suggested that if igneous rock, formed as a result of a prehistoric volcanic eruption, contanied a small amount of uranium it would steadily decay, leaving less uranium and depositing more stable *Pb*-206. By measuring the ratio of uranium to lead in rock samples, a rather exact time can be determined for the origin of the geological deposits. Uranium dating measures times of the order of millions of years.

Since all plants use CO_2 from the atmosphere for growth, a portion of the carbon in plants is radioactive *C*-14 and the plants are slightly radioactive. When a plant dies, no additional *C*-14 is taken in, and that within the plant body begins to decay without being replaced. Measurement of the relative amounts of *C*-14 and *C*-12 in an organic archeological sample provides a sensitive method of dating.

EXAMPLE. *A carbon specimen found in a cave contained 1/8 as much* C^{14} *as an equal amount of carbon in living matter. Calculate the approximate age of the specimen. Half-life period of* C^{14} *is* 5568 *years.*

SOL. Here, $$T_{1/2} = 5568 \text{ years}; \; T_{1/2} = \frac{0.6931}{\lambda}$$

or $$\lambda = \frac{0.6931}{_{1/2}} \quad \frac{0.6931}{5568} \text{ year}^{-1}$$

We have, $$N = N_0 e^{-\lambda t} \text{ or } \frac{N}{N_0} = e^{-\lambda t} \text{ or } \frac{1}{8} = e^{-\lambda t}$$

or $$\lambda t = \log_e 8$$

∴ $$t = \frac{\log_e 8}{\lambda} = \frac{(2.3026 \log_{10} 8) \times 5568}{0.6931}$$

$$= 16710 \text{ years.}$$

31.36 Biological Effects of Nuclear Radiations

The fact that nuclear radiation such as γ-rays, α-particles, β-particles, and neutrons can cause damage to the human body is well known.

The harmful effects of nuclear radiations appear to be due to the ionization or excitation of atoms in living cells by the Compton effect, bremsstrahlung, photoelectric effect, and so on. Some of the cell constituents are altered or destroyed by ionisation and some of the products formed may act as poisons. Examples of damages are the breaking up of chromosomes, swelling of the nucleus of a cell or of the entire cell; changes in the permeability of cell membrane, and destruction of cells.

The biological effects can be divided into three groups :

(*i*) Short-term recoverable effects;

(*ii*) Long-term irrecoverable effects; and

(*iii*) The genetic effect.

Groups (*i*) and (*ii*) are limited to the individuals who have actually received the radiation while the effects in group (*iii*) only appear in later generations.

When excessive doses are absorbed, the first noticeable disorder is a drop in the white blood cell count, which becomes evident in the first few hours after exposure. This is followed by a sickness pattern of diarrhoea,vomiting and fever which is now called radiation sickness. Recovery is possible from small acute doses. Large doses are lethal within a few weeks. Smaller doses produce short-term effects, such as skin disorders and loss of hair, which are generally recoverable. More serious is the damage done to the bone marrow and to other cells which may not have the ability to repair the damage. This leads to leukaemia and to the production of cancerous cells and finally malignant tumours.

The effects of the third type of damage only appear in the future generations of those irradiated. Experiments on animals show that the genetic effects to be expected from breeding from radiation-dosed human individuals include an increase in mental deficiencies, an increase in the number of monsters born and a general deterioration of the species in quality and population number.

The S.I. unit of radiation dosage is the *gray*. 1 *Gy* corresponds to 1 *J* of energy absorbed per kilogram of target material. The gray is a large unit and the *rad*, equal to 0.01 *Gy*, is more widely used. *One rad is defined as the absorbed dose of any nuclear radiation that results in the absorption of* 0.01 *J of energy per kilogram of absorbing material.* Since the absorbed dose required to produce a certain effect may be different for various types of radiation, this difference is expressed by the relative biological effectiveness (*RBE*) of the particular nuclear radiation. Finally, a dose unit for biological effect is the *rem* which is the dose in rads multiplied by the *RBE*. The *RBE* of 250 *keV* X-rays is taken as 1, and a dosage of 1 rad of such X-rays to a person is considered 1 *rem* (rad equivalent man). The *RBE*s of other X-rays and of gamma and beta rays are close to 1. But the *RBE* of fast neutrons is about 10 and that of 1-MeV alpha particles is about 25; 1 rad absorbed from such alpha particles therefore means a dosage of about 25 rem.

31.37 Non-Conservation of Parity in Beta Decay

Non-conservation of parity by neutrino. Prior to 1957, parity was assumed to remain conserved in all nuclear interactions, implying that the mirror image of an object or process is not distinguishable from the object or the process itself. In 1957, *Lee and Yang* hinted at the non-conservation of parity in β-decay.

Consider the mirror image of a neutrino. Let the neutrino travel toward a mirror with momentum $\vec{p}_\nu$. Its spin $\vec{s}_\nu$ is directed opposite to $\vec{p}_\nu$ (Fig. 31.27). Clearly, the mirror image of this neutrino is a different particle — the anti-neutrino — because in the mirror image, the vectors $\vec{s}_\nu$ and $\vec{p}_\nu$ are in the same direction. Thus, the motion of the neutrino as given by the mirror image is not a possible motion for it; it refers to a different particle viz., an antineutrino. This is known as *non-conservation of parity* by neutrino.

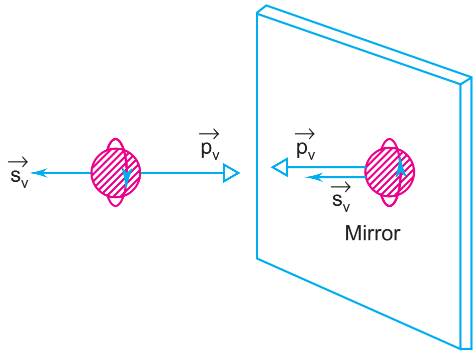


Fig. 31.27

Experimental verification of non-conservation of parity in β-decay

The violation of parity in the process of β-decay was experimentally demonstrated by *Wu* and coworkers.

Experimental arrangement. Fig. 31.28 shows the experimental set up used by *Wu et al.,* in their demonstration of the asymmetric β-particle distribution from polarized cobalt nuclei.

- The source consists of a thin layer of Co^{60} grown as a crystalline layer on top of good single crystal of cerium-magnesium-nitrate. Cobalt-60 emits an electron of energy 0.312 Mev in over 99% of its decays and two other electrons in the remaining decays. Several γ-rays are also emitted. Polarised samples of Co^{60} nuclei are prepared by applying the magnetic fields at low temperatures. The sample is cooled to 4.2 K by liquid helium and further cooled to below 0.01 K by adiabatic demagnetization.

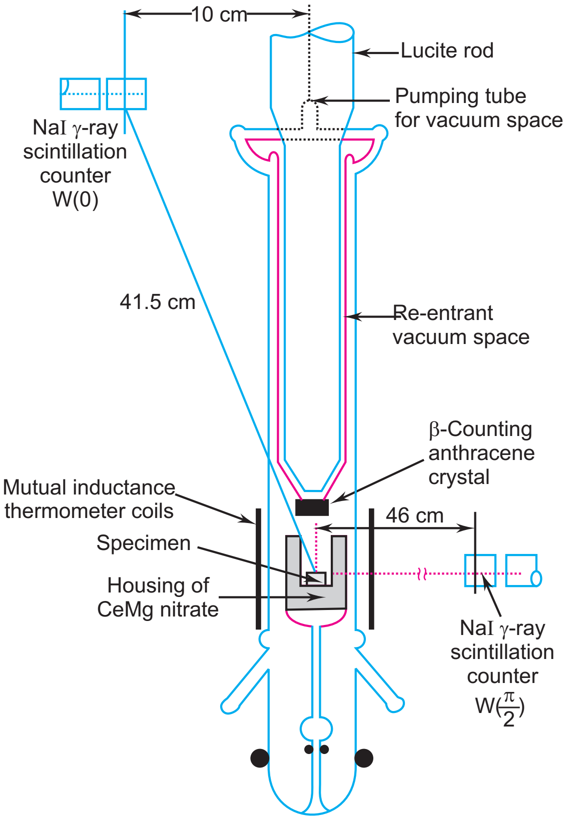

10 cm
Lucite rod
Pumping tube for vacuum space
NaI γ-ray scintillation counter W(0)
41.5 cm
Re-entrant vacuum space
β-Counting anthracene crystal
Mutual inductance thermometer coils
46 cm
Specimen
Housing of CeMg nitrate
NaI γ-ray scintillation counter $W(\frac{\pi}{2})$

Fig. 31.28

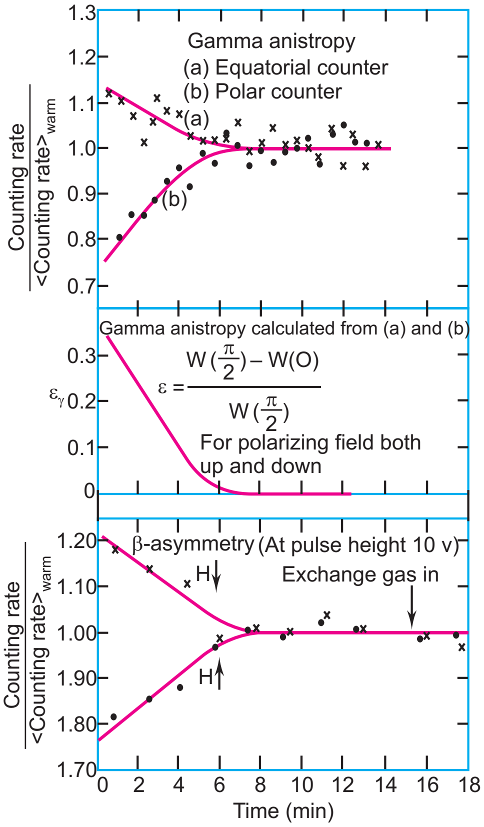


Fig. 31.29

- The beta-particles emitted by the polarized ^{60}Co nuclei are detected by an anthracene crystal mounted 2 cm above the source. The space between them is a vacuum. The light flashes in the scintillator, produced by the beta-rays, are transmitted through a lucite light pipe to a photomultiplier tube mounted outside the cryostat.
- To measure the extent of the polarization of the Co^{60} nuclei, two NaI γ-rays scintillation counters are used, one in the equatorial plane and one near the polar position. The observed γ-ray anisotropy provided a measure of the polarization and hence of the temperature of the sample.

Experiment. After the material in the cryostat is cooled to about 0.01 K, the magnet used in the cooling process is turned off. A vertical solenoid is raised around the lower part of the cryostat within 20 sec after the demagnetization so that the warming starts. The counting is started.

The results of the experiment. Fig. 31.29 shows the results of the experiment.

(*i*) Gamma-ray anisotropy. The *time scale is actually a temperature scale measuring time from the instant when cooling is stopped.* The curve labelled "gamma-anisotropy" is a measure of the extent of polarization. After about 8 min the nuclei have become sufficiently warm so that their thermal motion causes the alignment to become random, and the γ-anisotropy disappears.

(*ii*) Beta asymmetry. The curve labelled "β asymmetry" is the significant one. It indicates the number of electrons emerging in the direction of the magnetic field (the direction in which the nuclei are aligned), and the number emerging in the opposite direction. It is evident from the figure that more electrons are emitted in the direction opposite that of the magnetic field, *i.e.,* in the direction opposite to that in which the nuclei are aligned. The electrons are emitted in a preferred direction, the principle of right-left symmetry is violated, and parity is not conserved. The size of the effect is remarkable. The electron intensity in one direction along the axis of rotation was found to be 40% greater than in the opposite direction. There was no doubt about the result.

Review Question.

Define Parity. Explain the violation of parity conservation during β-decay process. How has this been experimentlly verified? (Meerut 1988)

EXERCISE

1. Explain the terms, decay constant, half-life and average life as applied to a radioactive substance. Find the relation between them.

2. Explain "half-period'' of a radioactive element and derive an expression for it. Calculate the decay constant of radium given its half-period as 1590 years. [**Ans.** 4.358×10^{-4} *year*$^{-1}$]

3. Give the theory of successive disintegration of radioactive substances. Explain what is radioactive equilibrium.

4. Mention the properties of α, β and γ-rays. Describe Curie's experiment to show the constituent radiations.

5. Explain Rutherford and Soddy's theory of radioactive decay. How is the half-life period defined?

The half-life period of uranium is 5000 million years. In how many years will 1 gm of the element be reduced to 1 centigram ?

6. Describe experiments to determine the charge and mass of α-particles. How is α-emission explained?

7. Give an account of experiments to determine the range and energy of α-particles accurately.

8. Describe an experiment for measuring the variation in the value of e / m with velocity in the case of β-particles.

9. Describe experiments for measuring the velocity and energy of β-particles.

10. Give the origin of the β-ray line and continuous spectrum. Outline the theory of β-disintegration.

11. Describe experiments to determine the wavelength of γ-rays.

12. If a radioactive element disintegrates for a period of time equal to its average life, what fraction of the original amount remains and what fraction will have disintegrated ? [**Ans.** 0.368; 0.632]

CHAPTER 32

RADIATION HAZARDS

AT A GLANCE

32.1	Radiation Hazards
32.2	Radiation Levels for Safety
32.3	Radiation Protection Methods
32.4	Nuclear Disasters
32.5	Nuclear Waste Disposal
32.6	Radiation Dosimeters

32.1 Radiation Hazards

There are two major sources of radiation contaminating the biosphere.

(*i*) People receive some natural or background radiation exposure each day from the sun, radioactive elements in the soil and rocks, household appliances (like television sets and microwave ovens), and medical and dental *x*-rays. These levels of natural and background radiation is normal. At sea level the average radiation due to natural background is about 100 millirad per year.

(*ii*) The second source of environmental contamination is man-made. This includes nuclear reactors, materials released through accidents or through the testing of nuclear devices underground or in the atmosphere. The radioactivity spread as a result of the latter is much higher than that due to the former source.

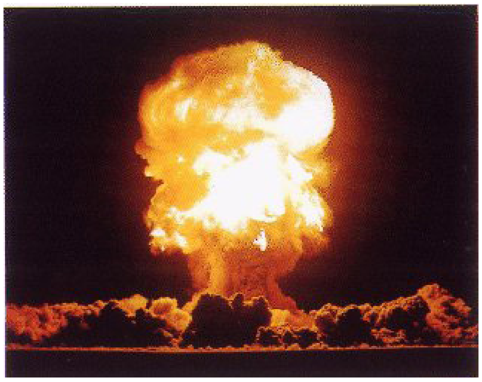
Radioactive materials-if handled improperly-or radiation accidentally released into the environment, can be dangerous because of the harmful effects of certain types of radiation on the body. The longer a person is exposed to radiation and the closer the person is to the radiation, the greater the risk.

32.2 Radiation Levels for Safety

The International Commission on Radiological Protection (ICRP) has recommended certain maximum permissible exposure (MPE) to radiation that is believed to result in no appreciable injury to a person. The ICRP requires that the maximum cumulative dose should not exceed 5 (N–18) REMS where N is the age of the person in years. The maximum short term dose should not exceed 3 REMS in three months a year, once a lifetime.

32.3 Radiation Protection Methods

The radiation exposure can be reduced to the safe limits by one or more of the following processes: by decreasing the time spent near a source, by allowing the source to decay for some time before approaching it, by increasing the distance from the source and by absorbing the radiation in shielding material. Other precautions include use of tongs and refraining from eating while using radioisotopes. Since persons receiving high or even lethal doses of radiation cannot sense it in any way, counters or dosimeters are used to determine if the prescribed maximum levels are exceeded. To cover individual workers in the range 10 milli REM to 1000 REM film badges are provided.



Nuclear Disasters (Chernobyl) Occurred in April 1986

Protection can also be afforded by increasing resistance to radiation through drugs.

32.4 Nuclear Disasters

The first major accident in the history of commercial nuclear energy in the U.S.A., occurred at Three Mile Island on March 28,1979.

The world's worst nuclear power accident occurred at Chernobyl in the former USSR (now Ukraine) on 26 April 1986. The reactor was used for electricity generation in Ukraine. The type of reactor used at Chernobyl is RBMK reactor. The Chernobyl nuclear power plant had 4 reactors. The chain reaction in reactor number 4 became out of control creating explosions. The radiation released from Chernobyl was 200 times more than that of the combined releases of the atom bombs that annihilated Hiroshima and Nagasaki in 1945. 260,000 square kilometres of territory in Ukraine, Russia and Belarus still have more than one curie per square kilometre of contamination with cesium 137.

A Vitrification Experiment for the Study of Nuclear Waste Disposal.

32.5 Nuclear Waste Disposal

The disposal of nuclear wastes is one of the formidable problems of nuclear energy projects. We have to think about the methods for disposal of nuclear wastes produced by reactors. Disposals in outer space, polar areas or in the sea bed are banned as they are unsafe. Studies are underway to bury these in the stable geological structures of bedded salt deposits. These days the waste containing radioactive materials is solidified after evaporation and is packed in metal canisters. The metal canisters are then stored away in big underground steel casks, each of about 35 tons capacity. The heat generated is removed by convection currents of air or chilled water.

The latest approach being adopted for disposal of nuclear wastes is to incorporate nuclear wastes in a boro *Silicate* glass which is similar to *pyrex*. The liquid wastes are converted into solid and added to normal ingredients for glass in the ratio of 1:3. The mixture is then heated till it melts and cast into blocks for deep *burial* into Earth. It may be noted that right location is easy to specify but very difficult to locate. We require that place which is stable geologically with no likelihood of *earthquakes,* no nearby population centres, a type of rock which does not disintegrate with heat and radiations but is soft enough for drilling etc. Also it should have almost no ground water which may be contaminated with radiations, etc.

32.6 Radiation Dosimeters

- A radiation dosimeter is a device used to detect the levels of expsoure to ionizing radiation. Dosimeters are frequently used in environments where exposure to radiation may occur, such as nuclear power plants and medical imaging facilities.

32.6.1 Thermoluminescence

Some materials, upon absorption of radiation, retain part of the absorbed energy in metastable states. When this energy is subsequently released in the form of ultraviolet, visible or infrared light, the phenomenon is called *luminescence.* Phosphorescence occurs with a time delay exceeding 10^{-8} s between stimulation and the emission of light. The process of phosphorescence can be accelerated with a suitable excitation. If the exciting agent is heat, the phenomenon is known as *thermoluminescence.* The material is called a thermoluminescent material, or a *TLD*.

The highly energetic secondary charged particles, usually electrons, that are produced in the primary interactions of photons with matter are mainly responsible for the photon energy deposition in matter. In a crystalline solid, these secondary charged particles release numerous low energy free electrons and holes through ionizations of atoms and ions. The free electrons and holes thus produced will either recombine or become trapped in an electron or hole trap, respectively, somewhere in the crystal.

The traps can be intrinsic or can be introduced in the crystal in the form of lattice imperfections consisting of vacancies or impurities. Two types of trap are known in general: storage traps and recombination centres.

- A storage trap merely traps free charge carriers and releases them during the subsequent heating, resulting in the thermoluminescence process.
- A charge carrier released from a storage trap may recombine with a trapped charge carrier of opposite sign in a recombination centre (luminescence centre). The recombination energy is at least partially emitted in the form of ultraviolet, visible or infrared light that can be measured with photomultiplier tubes (PMTs).

32.6.2 Thermoluminescent dosimeter

Principle. Thermoluminescent dosimeters, or TLDs, use thermoluminescent materials to detect radiation. These materials trap electrons. To read the dosage, the TLD is heated, which causes it to generate light that can be read by electronic equipment to determine the level of dosage.

Construction. Fig. 32.1 shows a schematic diagram of a TLD reader.

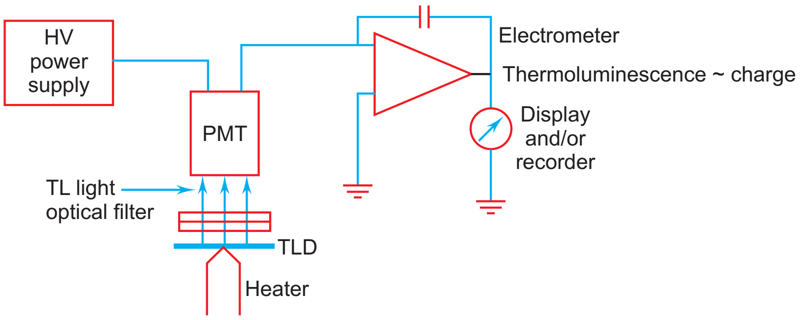

HV power supply
PMT
Electrometer
Thermoluminescence ~ charge
Display and/or recorder
TL light
optical filter
TLD
Heater

Fig. 32.1

- A **thermoluminescent dosimeter,** or ***TLD*,** is a type of radiation dosimeter. A *TLD* measures ionizing radiation exposure by measuring the amount of visible light emitted from a crystal in the detector when the crystal is heated. The amount of light emitted is dependent upon the radiation exposure.
- A basic *TLD* reader system consists of a planchet for placing and heating the *TLD*, a *PMT* to detect the thermoluminescence light emission and convert it into an electrical signal linearly proportional to the detected photon fluence and an electrometer for recording the PMT signal as a charge or current.
- The two most common types of TLDs are calcium fluoride and lithium fluoride, with one or more impurities to produce trap states for energetic electrons.

Energy Level Diagram of the *TL* Process

A phenomenological model of the thermoluminescence mechanism is provided in terms of the band model for solids (Fig. 32.2).

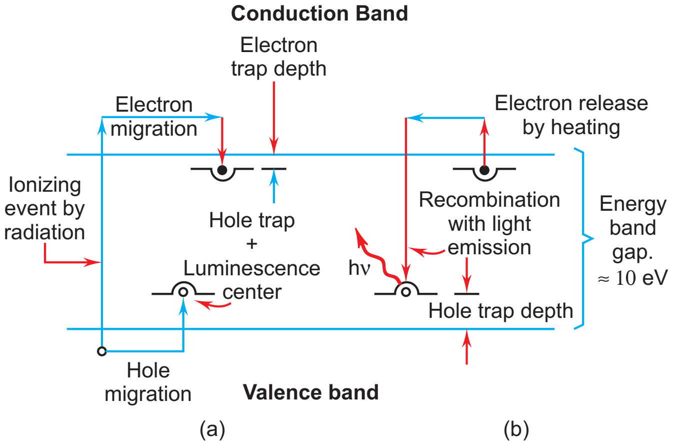

Conduction Band
Electron trap depth
Electron migration
Electron release by heating
Ionizing event by radiation
Hole trap + Luminescence center
Recombination with light emission
hν
Energy band gap. ≈ 10 eV
Hole trap depth
Hole migration
Valence band
(a)
(b)

Fig. 32.2

Fig. 32.2 (*a*) shows *ionization by radiation, and trapping of electrons and holes.*

Fig. 32.2 (*b*) shows *heating to release electrons, allowing luminescence production.*

The storage traps and recombination centres, each type characterized with an activation energy (trap depth) that depends on the crystalline solid and the nature of the trap, are located in the energy gap between the valence band and the conduction band. The states just below the conduction band represent electron traps, the states just above the valence band are hole traps. The trapping levels are empty before irradiation (*i.e.,* the hole traps contain electrons and the electron traps do not).

During irradiation the secondary charged particles lift electrons into the conduction band either from the valence band (leaving a free hole in the valence band) or from an empty hole trap (filling the hole trap).

- As the radiation interacts with the crystal it causes electrons in the crystal's atoms to jump to higher energy states, where they stay trapped due to intentionally introduced impurities (usually manganese or magnesium) in the crystal, until heated. Heating the crystal causes the electrons to drop back to their ground state, releasing a photon of energy equal to the energy difference between the trap state and the ground state.

32.6.3 Pocket Dosimeter

Pocket dosimeter is used to provide the wearer with an immediate reading of his or her exposure to *x*-rays and gamma rays. It is commonly worn in the pocket.

Construction. Fig. 32.3 shows pocket dosimeter.

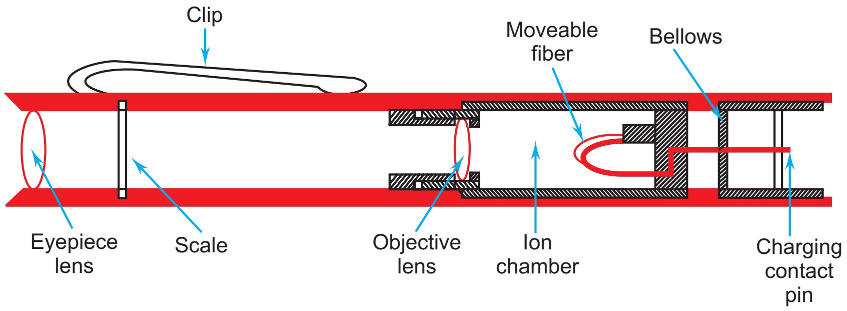


Fig. 32.3

It consists of a tiny ionization chamber that contains a small quartz fiber. The chamber and fiber are contained in a metal tube with a lens that allows the fiber to be viewed.

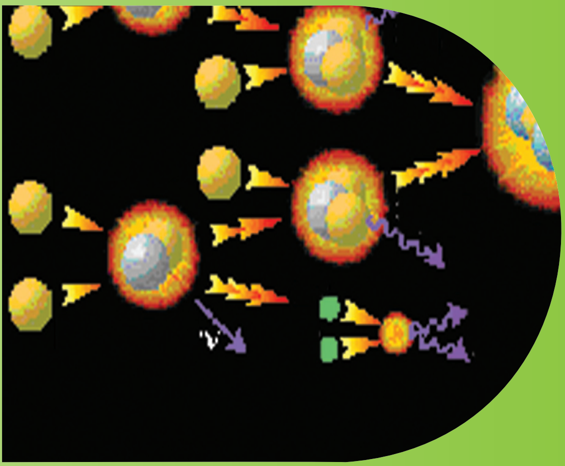
When exposed to radiation, the quartz fiber is deflected. When viewed through the eyepiece lens, this deflection can be easily seen against a scale. The scale is graduated in units of exposure. Typical industrial radiography pocket dosimeters have a full scale reading of 200 milliroentgens.

Digital Electronic Dosimeter

Another type of pocket dosimeter is the Digital Electronic Dosimeter. These dosimeters record dose information and dose rate. These dosimeters most often use Geiger-Müller counters. The output of the radiation detector is collected. When a predetermined exposure has been reached, the collected charge is discharged to trigger an electronic counter. The counter then displays the accumulated exposure and dose rate in digital form.

CHAPTER

33



Nuclear Reactions

AT A GLANCE	
33.1	Neutron Thermalization
33.2	Centre-of-Mass Coordinate System

33.1 Neutron Thermalization

Consider a particle of mass m_1 and initial velocity v_1 that collides head-on with another particle of mass m_2 at rest (Fig. 33.1). After collision they move with velocities v_1' and v_2'. The collision is elastic.

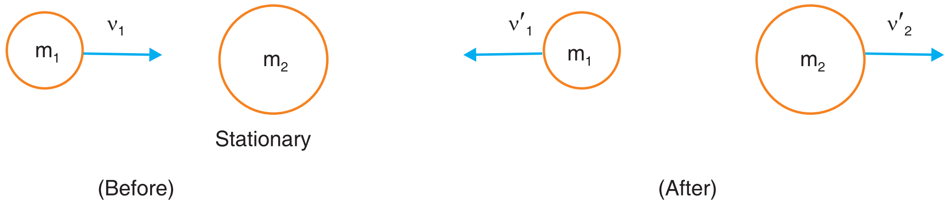


Fig. 33.1

By the law of conservation of linear momentum,

$$m_1 v_1 = m_2 v_2' - m_1 v_1'$$

or $$m_1 (v_1 + v_1') = m_2 v_2' \quad ...(1)$$

Since the collision is elastic, kinetic energy is also conserved.

$$\frac{1}{2} m_1 v_1^2 = \frac{1}{2} m_1 v_1'^2 + \frac{1}{2} m_2 v_2'^2$$

or $$m_1 (v_1^2 - v_1'^2) = m_2 v_2'^2 \quad ...(2)$$

Dividing Eq. (2) by Eq. (1), we obtain

$$v_1 - v_1' = v_2'$$

or $$v_1' = v_1 - v_2' \quad ...(3)$$

Substituting this value of v_1' in Eq. (1), we get

$$v_2' = \frac{2m_1}{(m_1+m_2)}v_1 \qquad \text{...(4)}$$

Let K_1 be the initial kinetic energy of the particle of mass m_1 and K_2' the kinetic energy lost by it in the collision. Then,

$$\frac{K_2'}{K_1} = \frac{\frac{1}{2}m_2 v_2'^2}{\frac{1}{2}m_1 v_1^2} = \frac{4m_1 m_2}{(m_1+m_2)^2} = \frac{4(m_2/m_1)}{[1+(m_2/m_1)]^2}$$

If $m_1 = m_2$, then $K_2' = K_1$ and the energy transfer is a maximum. It means that *all* the energy of m_1 is given to m_2. The neutron may transfer its entire kinetic energy to a proton on collison with it, since the mass $m_p \approx m_n$.

The above considerations show that in order to "slow down" the fast neutrons in a nuclear reactor they must be made to collide with the stationary targets (nuclei) of nearly the same mass as of the neutrons themselves ($m_1 = m_2$). That is why paraffin (which is rich in hydrogen whose nucleus 'proton' has nearly the same mass as a neutron) is a very good moderator.

33.2 Centre-of-Mass Coordinate System

When the law of conservation of energy and the law of conservation of momentum were written to solve the nuclear collision discussed in Section 33.1, it was implicit that the velocities in the various expressions were the values in the *laboratory coordinate system*. In this *L-system*, it was assumed that the target nucleus was at rest before the collision and that only the bombarding particle was in motion. The treatment of nuclear collisions is easier if the centre of mass of the particles is taken as the reference system. In the *centre-of-mass* or *C-system*, the centre of mass of the incident particle and the target nucleus is considered to be at rest and both particles are approaching it.

The relationship between the *L* and *C-systems* is shown in Fig. 33.2. In the *L-system*, the observer views the collision occurring as shown in Fig. 33.2 (*a*). A particle of mass m_1 and velocity v is incident upon a stationary particle of mass m_2. In the *C*-system, the observer views the collision occuring as shown in Fig. 33.2 (*b*). To an observer located at the centre of mass, the particles have equal and opposite momenta. The velocity V of the centre of mass is defined by the condition

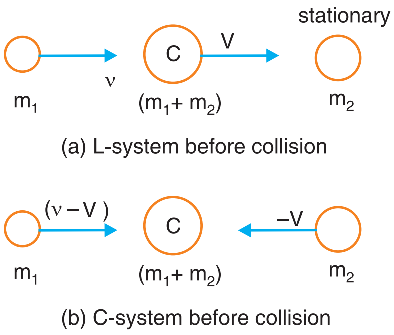


Fig. 33.2

$$m_1(v-V) = m_2 V$$

or

$$V = \left(\frac{m_1}{m_1+m_2}\right)v \qquad \text{...(1)}$$

In the *L*-system, the total kinetic energy is that of the incident particle only.

$$K_{lab} = \frac{1}{2}m_1 v^2 \qquad \text{...(2)}$$

In the *C*-system, m_1 has a velocity of $(v-V)$ towards the right and m_2 has a velocity of V towards the left. Therefore, the total kinetic energy in this reference frame is

$$K_{cm} = \frac{1}{2}m_1(v-V)^2 + \frac{1}{2}m_2 V^2$$

$$= \frac{1}{2}m_1 v^2 - \frac{1}{2}(m_1+m_2)V^2$$

$$= K_{lab} - \frac{1}{2}(m_1+m_2)V^2$$

$$\therefore \qquad K_{lab} = K_{cm} + \frac{1}{2}(m_1+m_2)V^2 \qquad ...(3)$$

$$\begin{matrix}\text{Total K.E. in} \\ \text{lab system}\end{matrix} = \begin{matrix}\text{Total K.E. in} \\ \text{cm system}\end{matrix} + \begin{matrix}\text{K.E. associated} \\ \text{with cm}\end{matrix}$$

When the particles collide, the maximum amount of kinetic energy that can be converted to excitation energy of the resulting compound nucleus while still conserving momentum is K_{cm}.

EXAMPLE 1. *Find the excitation energy of* ${}_{92}U^{236}$ *nucleus when* U^{235} *absorbs a thermal neutron of energy E* = 0.025 *MeV. Mass of* U^{235} = 235.0439 *u and those of* U^{236} *and Neutron are* 236.0457 *u and* 1.0087 *u respectively.* (*Osmania University, Oct.* 93)

SOL. Mass of U^{235}= 235.0439 *u*

Mass of neutron = 1.0087 *u*

Total mass of target and projectile = 235.0439 + 1.0087 = 236.0526 u

Mass of U^{236} = 236.0457 *u*

Mass defect Δm = (236.0526 – 236.0457) = 0.0069 *u*

The equivalent energy = 0.0069 × 931 = 6.424 MeV.

The excitation energy = Average binding energy per nucleon + KE of captured neutron

$$\left.\begin{matrix}\text{Average binding energy} \\ \text{per nucleon of } U^{236}\end{matrix}\right\} = \frac{6.424}{236} = 0.02723\,\text{MeV}$$

KE of projectile neutron = 0.025 MeV

Excitation energy = (0.02723 + 0.025) = 0.05223 MeV.

EXAMPLE 2. *The precise masses in the reaction*

$$H^1 + F^{19} \to He^4 + O^{16}$$

have been determined by mass spectrometer and are m (*H*) = 1.007825 *u; m* (*He*) = 4.002603 *u; m* (F^{19}) = 18.998405 *u; m* (O^{16}) = 15.994915 *u. Determine the Q and the nature of the reaction.*

(*Nagpur University, Oct.* 92)

SOL.

$$H^1 + F^{19} \to He^4 + O^{16} + Q$$

$$1.007825 + 18.998405 \to 4.002603 + 15.994915 + Q$$

$$Q = (+\,20.00623 - 19.997518)u = +\,0.008712\text{ u}$$

$$= +\,0.008712 \times 931\text{ MeV} = +\,8.11\text{ MeV}$$

As Q is + ve, the reaction is exoergic.

CHAPTER 34

ARTIFICIAL TRANSMUTATION OF ELEMENTS

AT A GLANCE	
34.1	The Discovery of Artificial Transmutation
34.2	Bohr's Theory of Nuclear Disintegration
34.3	The Q-Value Equation for a Nuclear Reaction
34.4	Nuclear Reactions
34.5	Energy Balance in Nuclear Reactions and the Q-value
34.6	Threshold Energy of an Endoergic Reaction
34.7	Nuclear Transmutations
34.8	The Scattering Cross-Section
34.9	Discovery
34.10	Preparation of Radioelements
34.11	Applications of Radioisotopes
34.12	The Discovery of the Neutron
34.13	Basic Properties of the Neutron
34.14	Classification of Neutrons
34.15	Neutron Sources
34.16	Neutron Detection
34.17	Neutron Collimator

34.1 The Discovery of Artificial Transmutation

Rutherford's experiment. Artificial transmutation is the conversion of one element into another provoked by artificial means. The apparatus with which Rutherford demonstrated the first artificial transmutation is shown in Fig. 34.1 (*a*). It consists of a chamber *C* one side of which has an opening covered by a thin sheet of silver foil *F*. A zinc sulphide screen *S* is placed close to *F* and the scintillations on it are observed by a microscope *M*. The chamber can be filled with various gases through the side tubes *TT*. The source of α-particles (*RaC'*) was placed at *R* on a small disc. The distance of *R* from *F* can be varied by means of a long rod attached to the disc at one end and manipulated from outside the chamber at the other. The range of the α-particle used was known to be about 7 cm in air, corresponding to an energy of about 8 MeV. The effects produced by these α-particles on the various gases can be studied by observing the scintillations on *S*.

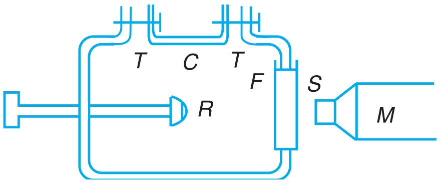


Fig. 34.1 (*a*)

When the chamber was filled with oxygen or CO_2 no scintillations were observed on the screen when the distance between *R* and *F* was greater than the range of the α-particles (7 cm). But when the chamber was filled with nitrogen, scintillations were observed on the screen even when the distance between *R* and *F* was as high as 40 cm. The α-particles themselves could not penetrate such a distance. Thus, obviously, scintillations were not caused by the α-particles but by particles emitted by the nitrogen atoms. Magnetic deflection experiments indicated that these particles were hydrogen nuclei or protons. On the basis of the Bohr's theory of the compound nucleus, the reaction can be written as

$${}_7N^{14} + {}_2He^4 \rightarrow ({}_9F^{18}) \rightarrow {}_8O^{17} + {}_1H^1$$

The reaction may be pictorially represented as in Fig. 34.1 (*b*). The new element formed was the rare isotope of oxygen of mass number 17. Thus, the first transmutation of nitrogen into oxygen was achieved and established.

Ernest Rutherford (1871–1937)

In 1919 at Manchester, he discovered that the nuclei of certain light elements, such as nitrogen, could be "disintegrated" by the impact of energetic alpha particles coming from some radioactive source, and that during this process fast protons were emitted.

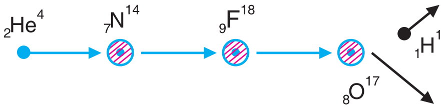


Fig. 34.1 (*b*)

34.2 Bohr's Theory of Nuclear Disintegration

Bohr assumed that the nucleus is somewhat like a liquid drop. When a projectile strikes the nucleus, the projectile is captured by the nucleus. The nucleus formed after the projectile is captured, is called a *compound nucleus*. The energy of the projectile is quickly distributed among the nucleons inside the compound nucleus. Therefore, the compound nucleus resembles a heated-up liquid drop and may be considered to exist in a variety of quasi-stationary excitation states. The compound nucleus persists in its excited state until a particular nucleon momentarily happens to have a large fraction of the excitation energy to escape from the compound nucleus, thus, leading to a disintegration. This process is similar to the slow evaporation of particles from the surface of a liquid drop.

A nuclear reaction may be represented by

$$a + X \rightarrow Y + b.$$

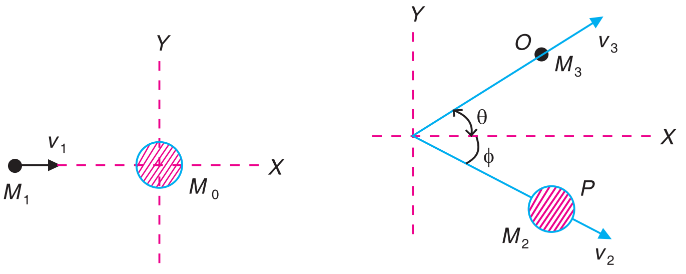
This formula signifies that the particle a interacts with the nucleus X to yield the nucleus Y and particle b. The above equation is usually abbreviated and is written as

$$X\,(a, b)\;Y.$$

For example, ${}_6C^{12}\,(d, n)\;{}_7N^{13}$ stands for a reaction between an incident deuteron ($d = {}_1H^2$) and a ${}_6C^{12}$ nucleus to produce a ${}_7N^{13}$ nucleus with the emission of a neutron.

34.3 The Q-Value Equation for a Nuclear Reaction

Let a particle of mass M_1 moving with velocity v_1 collide with a target nucleus of mass M_0 at rest (Fig. 34.2). After collision, the particle O of mass M_3 is emitted with velocity v_3 at an angle θ and the recoiling nucleus P of mass M_2 is emitted with a velocity v_2 at an angle ϕ. The conservation of linear momentum in the plane of paper yields the equations



(*a*) Before collision (*b*) After collision

Fig. 34.2

$$M_1 v_1 = M_3 v_3 \cos\theta + M_2 v_2 \cos\phi \qquad \text{...(1)}$$

$$0 = M_3 v_3 \sin\theta - M_2 v_2 \sin\phi \qquad \text{...(2)}$$

$$\therefore \quad M_2\, v_2 \cos\phi = M_1 v_1 - M_3 v_3 \cos\theta \qquad \text{...(3)}$$

$$M_2\, v_2 \sin\phi = M_3 v_3 \sin\theta \qquad \text{...(4)}$$

Squaring and adding, we get

$$M_2^2 v_2^2 = M_1^2 v_1^2 + M_3^2 v_3^2 - 2M_1\, M_3\, v_1\, v_3 \cos\theta \qquad \text{...(5)}$$

Using the kinetic energy relations,

$$E_{k_1} = \frac{1}{2} M_1 v_1^2,\; E_{k_2} = \frac{1}{2} M_2 v_2^2 \text{ and } E_{k_3} = \frac{1}{2} M_3 v_3^2$$

Substituting these values in Eq. (5), we have

$$2M_2\, E_{k_2} = 2M_1\, E_{k_1} + 2M_3\, E_{k_3} - 2\,(M_1\, M_3\, E_{k_1} E_{k_3})^{1/2}\, 2\cos\theta$$

or

$$E_{k_2} = \frac{M_1\, E_{k_1}}{M_2} + \frac{M_3\, E_{k_3}}{M_2} - \frac{2}{M_2}(M_1\, M_3\, E_{k_1}\, E_{k_3})^{1/2} \cos\theta \qquad \text{...(6)}$$

Let

$$Q = (E_{k_2} + E_{k_3})\; E_{k_1} \qquad \text{...(7)}$$

Substituting the value of E_{k_2} from Eq. (6) in Eq. (7), we have

$$Q = \frac{M_1 E_{k_1}}{M_2} + \frac{M_3}{M_2} E_{k_3} - \frac{2}{M_2}(M_1 M_3 E_{k_1} E_{k_3})^{1/2} \cos\theta + E_{k_3} - E_{k_1}$$

or
$$Q = E_{k_3}\left(1+\frac{M_3}{M_2}\right) - E_{k_1}\left(1-\frac{M_1}{M_2}\right) - \frac{2}{M_2}(M_1 M_3 E_{k_1} E_{k_3})^{1/2} \cos\theta \quad \text{...(8)}$$

If $\theta = 90°$, $\cos 90° = 0$. Hence, Eq. (8) reduces to

$$Q = E_{k_3}\left(1+\frac{M_3}{M_2}\right) - E_{k_1}\left(1-\frac{M_1}{M_2}\right) \quad \text{...(9)}$$

Threshold energy. *The threshold energy* (E_{th}) *is defined as the minimum kinetic energy of the incident particle which will initiate an endoergic reaction.*

We can calculate the threshold energy by using the centre of mass coordinate system. For any collision or reaction in centre of mass coordinate system, the linear momentum is always zero before and after the reaction. Let E'_{k_1} be the kinetic energy of the incident particle in centre of mass coordinate system. An endoergic reaction is possible if

$$E'_{k_1} \geq |Q| \quad \text{...(10)}$$

Let M_{red} be the reduced mass of incident particle and the target nucleus. Then,

$$M_{red} = \frac{M_1 M_0}{(M_0 + M_1)}$$

$$\therefore \quad E'_{k_1} = \frac{1}{2} M_{red} v_1^2 = \frac{1}{2} \frac{M_1 M_0}{(M_0 + M_1)} v_1^2 \quad \text{...(11)}$$

From Eqs. (10 and 11),

$$\frac{1}{2} \frac{M_1 M_0}{(M_0 + M_1)} v_1^2 \geq |Q| \text{ or } \frac{1}{2} M_1 v_1^2 \geq \left(\frac{M_0 + M_1}{M_0}\right) |Q|$$

$$\therefore \text{ Threshold energy} \quad E_{th} = \left(1 + \frac{M_1}{M_0}\right) |Q| \quad \text{...(12)}$$

34.4 Nuclear Reactions

Following are some of the main types of nuclear reactions.

***(i)* Elastic scattering.** In this case the incident particle strikes the target nucleus and leaves without loss of energy, but its direction may change.

EXAMPLE : Scattering of α-particles from a thin gold foil.

$${}_{79}Au^{197} + {}_2He^4 \rightarrow {}_{79}Au^{197} + {}_2He^4$$

The target nucleus remains unaffected.

***(ii)* Inelastic scattering.** In this case, the incident particle loses a part of its energy in exciting the target nucleus to a higher allowed energy level. The excited nucleus later decays to the ground state, radiating the excess energy in the form of a γ-ray photon. *Example* :

$${}_3Li^7 + {}_1H^1 \rightarrow {}_3Li^{7*} + {}_1H^1$$

$$_3Li^{7*} \rightarrow {_3Li^7} + \gamma.$$

(*iii*) Radiative Capture. Here the incident particle is captured by the target nucleus and a new nucleus is formed. The new nucleus, in general, has a considerable excess of energy and decays with the emission of one or more γ-ray photons. *Example* :

$$_6C^{12} + {_1H^1} \rightarrow {_7N^{13*}} \rightarrow {_7N^{13}} + \gamma$$

(*iv*) Disintegration. Here the incident particle is absorbed by the target nucleus and the ejected particle is a different one. The composition of the resultant nucleus is also different from the parent nucleus. An example is the disintegration of beryllium by α-particle producing neutrons.

$$_4Be^9 + {_2He^4} \rightarrow {_6C^{12}} + {_0n^1}$$

(*v*) Photodisintegration. When target materials are bombarded with radiations, the resulting compound nuclei are usually formed in excited states. These nuclei generally get rid of the excess excitation energy through neutron emission. For example,

$$_1H^2 + \gamma \rightarrow {_1H^1} + {_0n^1}$$

This requires a photon of energy 2.225 MeV.

Conservation Laws. The conservation laws may be stated and illustrated by reference to some specific nuclear reaction, say,

$$_5B^{10} + {_2He^4} \rightarrow ({_7N^{14}}) \rightarrow {_1H^1} + {_6C^{13}} + Q$$

(*i*) Conservation of charge : Total charge is conserved in every type of *nuclear reaction*. In B^{10} (α, *p*) C^{13} there are seven protons initially, also seven in the compound nucleus and in the products of the reaction.

(*ii*) Conservation of nucleons : *The total number of nucleons entering and leaving the reaction is constant.* In B^{10} (α, *p*) C^{13}, we find 14 nucleons at each stage of the reaction.

(*iii*) Conservation of mass-energy : *In nuclear reactions neither kinetic energy nor rest mass is conserved by itself. But their total is always conserved.*

(*iv*) Conservation of parity : *The net parity before the reaction must equal the net parity after the reaction.*

(*v*) *Linear momentum, angular momentum, spin and isotopic spin are the other physical quantities which are also conserved in a nuclear reaction.*

(*vi*) Quantities not conserved : *The most prominent physical characteristics which are not conserved in nuclear reactions are the magnetic dipole moments and the electric quadrupole moments of the reacting nuclei. These moments depend upon the internal distribution of mass, charge and current within the nuclei involved and are not subject to conservation laws.*

34.5 Energy Balance in Nuclear Reactions and the Q-value

In all nuclear reactions, the total sum of mass and energy is conserved. Thus taking the equation $A + B \rightarrow P + O$, the target is supposed to be at rest and let its mass be m_1. The projectile has a mass m_2 and $K.E. = E_2$.

The product nucleus has a mass m_3 and K.E. E_3. The outgoing particle has a mass m_4 and K.E. E_4. The equation representing the conservation of total energy is written as

$$m_1c^2 + m_2c^2 + E_2 = m_3c^2 + E_3 + m_4c^2 + E_4$$

Let $$Q = E_3 + E_4 - E_2 = (m_1 + m_2 - m_3 - m_4)c^2.$$

The quantity Q is called the energy balance or the Q-value of the reaction. If Q is positive, the reaction is *exothermic* (or *exoergic*). Now the K.E. of products of transmutation is greater than the K.E. of the reactants and energy is released in the process. If Q is negative, the reaction is *endothermic* (or *endoergic*). The equation for a nuclear reaction is written as

$$A + B \rightarrow P + O + Q$$

where Q may be positive or negative.

Example 1. *Consider the reaction*

$${}_{7}N^{14} + {}_{2}He^{4} \rightarrow {}_{8}O^{17} + {}_{1}H^{1} + Q$$

where Q denotes the energy absorbed or evolved during the nuclear reaction.

Sol. The atomic masses of the particles are : $N^{14} = 14.003074$, $He^{4} = 4.002604$, $O^{17} = 16.99913$ and $H^{1} = 1.007825u$.

Substituting these values in the equation,

$$14.003074 \quad + \quad 4.002604 \quad = \quad 16.99913 \quad + \quad 1.007825 \quad + \quad Q$$

$$\therefore \qquad Q \quad = \quad 18.005678 \quad - \quad 18.006955 \quad = \quad -0.001277u$$

According to Einstein's mass-energy relation, $1\ u = 931.3$ MeV

$\therefore$ The calculated value of $Q = -0.001277 \times 931.3 = -1.189$ MeV.

This is in close agreement with the experimental value (*i.e.*, –1.26 MeV). Incidentally, Einstein's mass-energy relation gets verified.

Example 2. *Determine the product nuclei and Q values in the following reactions : Al^{27} (d, α) and Mg^{25} (α, d). Masses of Al^{27}, Mg^{25}, α and d are 26.9901, 24.9936, 4.0039 and 2.0147 amu respectively. Comment on your results.*

Sol. (*i*) The nuclear reaction is

$${}_{13}Al^{27} + {}_{1}H^{2} \rightarrow {}_{12}Mg^{25} + {}_{2}He^{4} + Q$$

$$\therefore \quad Q = 26.9901 + 2.0147 - 24.9936 - 4.0039 = 0.0073 \text{ amu.}$$

$$i.e., \quad Q = 0.0073 \times 931.3 \text{ MeV} = 6.797 \text{ MeV.}$$

(*ii*) $${}_{12}Mg^{25} + {}_{2}He^{4} \rightarrow {}_{13}Al^{27} + {}_{1}H^{2} + Q$$

$$\therefore \quad Q = 24.9936 + 4.0039 - 26.9901 - 2.0147 = -0.0073 \text{ amu.}$$

$$i.e., \quad Q = -0.0073 \times 931.3 \text{ MeV} = -6.797 \text{ MeV.}$$

34.6 Threshold Energy of an Endoergic Reaction

In an endoergic reaction $-Q$ is the energy which is needed to excite the reaction. This energy is supplied by K.E. of the incoming particle. But all the K.E. of the projectile particle is not available for the nuclear reaction because part of the energy is used to give K.E. to the compound nucleus. Hence, for a reaction in which $-Q$ energy is to be absorbed, the incident particle should supply some energy in addition to $-Q$. Thus the minimum K.E. which the projectile should possess so that the nuclear reactions may take place is called the *threshold energy*.

Let m_i and v_i be the mass and initial velocity of the projectile. Let m_c and v_c be the mass and velocity of the compound nucleus. Applying conservation of momentum principle,

$$m_i v_i = m_c v_c \text{ or } v_c = m_i v_i / m_c.$$

$$\text{K.E. of compound nucleus} = \frac{1}{2} m_c v_c^2 = \frac{1}{2} m_c \left(\frac{m_i}{m_c}\right)^2 v_i^2$$

Energy available for the reaction $= \frac{1}{2}m_i\, v_i^2 - \frac{1}{2}m_c\, v_c^2 = \frac{1}{2}m_i\, v_i^2\left(1-\frac{m_i}{m_c}\right)$

But $m_c = m_i + m_t$ where m_t = mass of target nucleus.

$$\therefore \qquad \frac{1}{2}m_i v_i^2 - \frac{1}{2}m_c v_c^2 = \frac{1}{2}m_i v_i^2\left(\frac{m_t}{m_i + m_t}\right)$$

or

$$-Q = \frac{1}{2}m_i v_i^2\left(\frac{m_t}{m_i + m_t}\right)$$

$$\therefore \qquad \text{The threshold energy} = E_{th} = \frac{1}{2}m_i v_i^2 = -Q\left(\frac{m_i + m_t}{m_t}\right)$$

Example. *The Q value of the Na^{23} (n, α) F^{20} reaction is –5.4 MeV. Determine the threshold energy of the neutrons for this reaction.*

Sol. Here, $Q = -5.4$ MeV; $m_i = 1.008665u$, $m_t = 22.9898\, u$.

$$E_{th} = -Q\left(\frac{m_i + m_t}{m_t}\right) = -(-5.4)\left(\frac{1.008665 + 22.9898}{22.9898}\right)$$

$$= 5.635 \text{ MeV.}$$

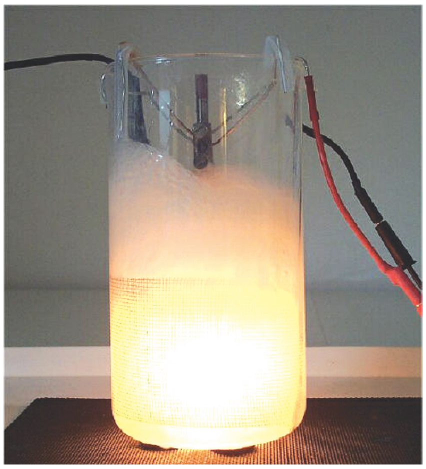
34.7 Nuclear Transmutations

Transmutations by α-particles. (*i*) (α, p) **reaction.** The elements are bombarded by α-particles and protons are ejected. Rutherford and Chadwick succeeded in disintegrating a number of elements from boron to potassium by α-particle bombardment. The equations are :

$${}_5B^{10} + {}_2He^4 \rightarrow {}_6C^{13} + {}_1H^1$$
$${}_7N^{14} + {}_2He^4 \rightarrow {}_8O^{17} + {}_1H^1$$
$${}_{19}K^{39} + {}_2He^4 \rightarrow {}_{20}Ca^{42} + {}_1H^1$$

In these reactions, the atomic number of the product nucleus increases by 1 and mass number by 3.

(*ii*) (α, n) **reaction.** When certain nuclei are bombarded by α-particles, neutrons are ejected. In this type of reaction, the atomic number increases by 2 and mass number by 3. It was this type of reaction that led to the discovery of neutrons.



Nuclear Transmutations.

Example.

$${}_4Be^9 + {}_2He^4 \rightarrow {}_6C^{13} \rightarrow {}_6C^{12} + {}_0n^1$$
$${}_5B^{11} + {}_2He^4 \rightarrow {}_7N^{15} \rightarrow {}_7N^{14} + {}_0n^1$$

Transmutation by protons. (*i*) (p, α) **reactions.** Cockcroft and Walton artificially accelerated the protons with their tension multiplier and used them as projectiles in nuclear reactions. When lithium was bombarded by protons having energies from 100 to 700 keV, lithium broke up into two α-particles after proton capture.

$${}_3Li^7 + {}_1H^1 \rightarrow 2\ {}_2He^4$$

There are many other disintegrations produced by the bombardment with protons and they are given below :

$${}_5B^{11} + {}_1H^1 \rightarrow {}_6C^{12} \rightarrow {}_4Be^8 + {}_2He^4$$

$${}_{11}Na^{23} + {}_1H^1 \rightarrow {}_{12}Mg^{24} \rightarrow {}_{10}Ne^{20} + {}_2He^4$$

(*ii*) (p, d) **reaction.** In some cases instead of α-particles, deuterons are also emitted.

$${}_4Be^9 + {}_1H^1 \rightarrow {}_5B^{10} \rightarrow {}_4Be^8 + {}_1H^2$$

(*iii*) (p, n) **reaction.** $${}_5B^{11} + {}_1H^1 \rightarrow {}_6C^{11} + {}_0n^1$$

(*iv*) (p, γ) **reaction.** $${}_3L^{17} + {}_1H^1 \rightarrow {}_4Be^8 + \gamma$$

Transmutation by deuterons. In many nuclear reactions, high energy deuterons are used as bombarding particles. The deuterons are accelerated in a cyclotron and they possess very high energy. Some of the nuclear reactions are :

(*i*) (d, α) **reactions.** $${}_3Li^6 + {}_1H^2 \rightarrow {}_2He^4 + {}_2He^4$$

$${}_8O^{16} + {}_1H^2 \rightarrow {}_7N^{14} + {}_2He^4$$

(*ii*) (d, p) **reactions.** $${}_6C^{12} + {}_1H^2 \rightarrow {}_6C^{13} + {}_1H^1$$

$${}_{11}Na^{23} + {}_1H^2 \rightarrow {}_{11}Na^{24} + {}_1H^1$$

(*iii*) (d, n) **reactions.** $${}_6C^{12} + {}_1H^2 \rightarrow {}_1H^2 \rightarrow {}_7N^{13} + {}_0n^1$$

$${}_4Be^9 + {}_1H^2 \rightarrow {}_5B^{10} + {}_0n^1$$

Transmutation by neutrons. Neutrons are very effective in producing transmutations. They are not affected by electrostatic fields because they do not possess any charge and are not repelled by the nuclei. Due to these reasons, neutrons have been found to be more penetrating into the nuclei as compared to α-particles, protons and deuterons. The disintegrations produced by neutrons leads to the emission of various particles like α-particles, protons, neutrons, photons and deuterons.

(*i*) (n, α) **reactions**

$${}_3Li^6 + {}_0n^1 \rightarrow ({}_3Li^7) \rightarrow {}_1H^3 + {}_2He^4$$

$${}_5B^{10} + {}_0n^1 \rightarrow ({}_5B^{11}) \rightarrow {}_3Li^7 + {}_2He^4$$

(*ii*) (n, p) **reactions**

$${}_7N^{14} + {}_0n^1 \rightarrow ({}_7N^{15}) \rightarrow {}_6C^{14} + {}_1H^1$$

$${}_{13}Al^{27} + {}_0n^1 \rightarrow ({}_{13}Al^{28}) \rightarrow {}_{12}Mg^{27} + {}_1H^1$$

(*iii*) (n, 2n) **reactions**

$${}_{13}Al^{27} + {}_0n^1 \rightarrow ({}_{13}Al^{28}) \rightarrow {}_{13}Al^{26} + 2\ {}_0n^1$$

(*iv*) (n, γ) **reactions**

$${}_{13}Al^{27} + {}_0n^1 \rightarrow ({}_{13}Al^{28}) \rightarrow {}_{13}Al^{28} + \gamma$$

$${}_{92}U^{238} + {}_0n^1 \rightarrow ({}_{92}U^{239}) \rightarrow {}_{92}U^{239} + \gamma$$

These reactions are called radiative captures. Such reactions usually produce artificially radioactive isotopes.

Disintegration by electrons. Electrons of very high energy obtained from betatrons cause the disintegration of a number of elements.

Example. *When beryllium target is bombarded with high energy electrons, neutron is ejected.*

$$_4Be^9 + {}_{-1}e^0 \rightarrow ({}_3Li^9) \rightarrow {}_3Li^8 + {}_0n^1$$

34.8 The Scattering Cross-Section

The properties of scattering interactions are usually expressed in terms of the *Scattering cross-section*. To define this term, we consider the following scattering experiment : A beam of particles (called *beam particles*) is directed on a scatterer consisting of *target particles*. As a result of *collisions* between beam and target particles, there are particles which emerge from the reactions (called *reaction products*) and these are detected in *particle detectors*. To describe this quantitatively, we suppose that the scatterer contains N_t target particles. The target is uniformly illuminated by a flux F_s (expressed as the number of beam particles per unit area per unit time arriving at the target) of beam particles. We suppose also that the scatterer is sufficiently small that the beam is negligibly attenuated in passing through it. Then, if there are δN_s scattering interactions per unit time which lead to detected particles, we define the scattering cross section $\delta\sigma$ as $\delta\sigma = \delta N_s / N_t F_s$. In the limit that the detectors subtend very small solid angles, as seen from the target, we define the *differential* scattering cross section $d\sigma$. When a single detector subtending a solid angle $\delta\Omega$ is used to define $\delta\sigma$, we may define the cross section per unit solid angle as

$$\frac{d\sigma}{d\Omega} = \lim_{\delta\Omega \to 0} \frac{\delta\sigma}{\delta\Omega}$$

The total scattering cross section σ is obtained by summing $\delta\sigma$ over all scattering events: $\sigma = \Sigma\delta\sigma$.

It is customary practice to associate a cross-section with each particular type of nuclear interaction. A scattering cross-section is used when dealing with nuclear scattering processes; an absorption cross-section for nuclear absorption processes; and a fission cross-section when studying nuclear collisions, leading to the fission of the target nucleus.

Determination of cross-section. Consider a slab of some material whose area is A and thickness is dx (Fig. 34.3). Its volume is Adx. If the target material contains n nuclei per unit volume, the total No. of nuclei in the slab = $nAdx$. Each nucleus has a cross-section of σ for some particular interaction, so that the aggregate cross-section of all the nuclei in the slab is $nA\,\sigma\,dx$. Let N be the No. of incident particles in a bombarding beam. Let dN be the No. of particles that interact with nuclei in the slab. Then,

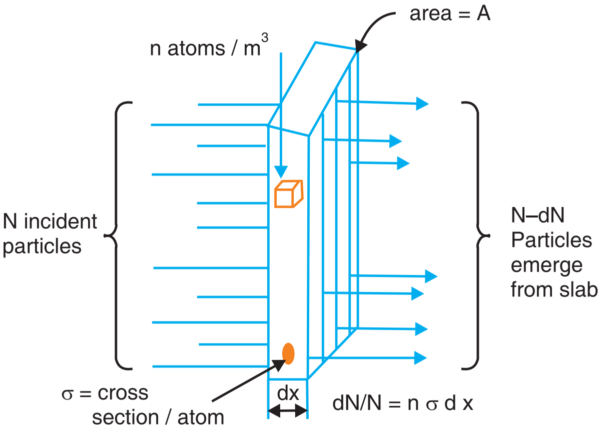


Fig. 34.3

$$\frac{dN}{N} = \frac{\text{Aggregate cross section}}{\text{target area}} = \frac{nA\sigma dx}{A} = n\sigma dx.$$

Now let us consider the same beam of particles incident on a slab of thickness x. Let N_0 be the number of incident particles. If each particle can interact only once, dN particles may be thought of as being removed from the beam in passing through the first dx of the slab. Hence

$$\frac{-dN}{N} = n\sigma dx \quad \text{or} \quad \int_{N_0}^{N} \frac{dN}{N} = -n\sigma \int_0^x dx$$

$$\text{or} \qquad \log_e N - \log_e N_0 = -n\sigma x$$

$$\text{or} \qquad N = N_0 e^{-n\sigma x}$$

The number of surviving particles N decreases exponentially with increasing slab thickness x. Here, σ is the cross-section per nucleus or the *microscopic cross-section.* n is the number of nuclei per unit volume. The product $n\sigma$ is called the *macroscopic cross-section and is denoted by* Σ. Σ *has the dimensions of* L^{-1}.

$$\text{Thus} \qquad N = N_0 e^{-\Sigma x}$$

The unit used for σ is an area of 10^{-28} m^2. This unit is called the *barn*.

If the incident flux (beam intensity) is I_0 and the flux after penetrating a distance x is I, then $I = I_0 e^{-\Sigma x}$ or

$$\log_e I = \log_e I_0 - \Sigma x$$

Σ can be found by measuring the beam intensity with and without targets. A plot of $\log_e I$ against x for several foil thicknesses gives Σ graphically and hence σ.

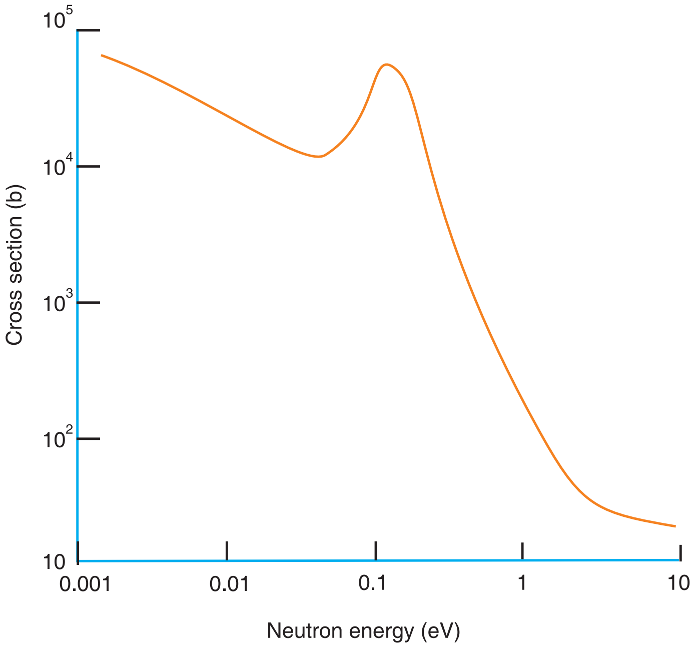


Fig. 34.4

The cross-sections for most nuclear reactions depend upon the energy of the incident particle. Fig. 34.4 shows how the neutron capture cross-section of $_{48}Cd^{113}$ varies with neutron energy. In this reaction, the absorption of a neutron is followed by the emission of a gamma ray : Cd^{113} (n, γ) Cd^{114}. The narrow peak at 0.176 eV is a resonance effect associated with an excited state in the Cd^{114} nucleus. The capture cross-sections of Cd^{113} for slow neutrons are so great that cadmium is widely used in control rods for nuclear reactors.

Macroscopic cross-section and mean free path. The mean free path (λ) of a particle in a material is the average distance it can travel in the material before interacting there.

$$l = \frac{\int_0^\infty x e^{-\Sigma x}\, dx}{\int_0^\infty e^{-\Sigma x}\, dx} = \frac{1}{\Sigma}$$

EXAMPLE. *The usefulness of cadmium in a nuclear reactor depends on the high thermal absorption cross-section of the 113 isotope. viz.,* $\sigma_\alpha\ (Cd^{113}) = 21000$ *barns. If the density of cadmium is* $8.7 \times 10^3\ kg\ m^{-3}$*, calculate the macroscopic cross-section of* Cd^{113} *and hence the thickness required to attenuate a neutron beam to 0.01% of its original intensity.*

SOL. Macroscopic cross section = $\Sigma = \sigma N_v$

where N_v = Number of nuclei per unit volume in the absorber.

Now unit volume of material of density ρ contains $N_A\rho/A$ nuclei, where N_A is the Avogadro constant and A is the atomic mass number of the absorber atoms.

Thus $$N_v = \frac{N_A}{A}\rho \quad \text{and} \quad \Sigma = \sigma\frac{N_A}{A}\rho$$

∴ $$\Sigma = \frac{(6.02\times10^{26})(8.7\times10^{3})(21000\times10^{-28})}{113} = 97000\,\text{m}^{-1}$$

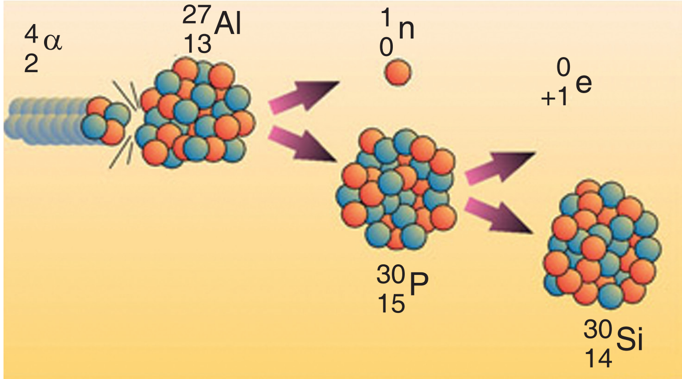
From $I = I_0e^{-\Sigma x}$ we have

$$\Sigma x = \log_e\left(\frac{I_0}{I}\right) = 2.3\log_{10}10000 = 2.3\times4 = 9.2$$

∴ $$x = \frac{9.2}{97000}\text{m} = 95\times10^{-6}\text{ m}.$$

ARTIFICIAL RADIOACTIVITY

34.9 Discovery



In 1934, Joliot-Curie Succeeded in Making the First Artificial Radioactive Nuclide.

Artificial or induced radioactivity was discovered by Curie and Joliot in 1934. They were studying the disintegration of light elements by α-particles. When boron and aluminium were bombarded with α-particles, the target continued to emit radiations even after the source of α-particles had been removed. With the help of magnetic deflection experiments and ionisation measurements it was found that the radiations consisted of particles having a positive charge and mass equal to electrons. These particles are known as positrons. Curie and Joliot explained the phenomenon as follows: When the above elements are bombarded by α-particles, an unstable nucleus was formed and this nucleus disintegrated spontaneously. In the case of boron, the nuclear reactions are

$${}_5B^{10} + {}_2He^4 \rightarrow {}_7N^{13*} + {}_0n^1$$

The nitrogen is radioactive and decays with a half-life of 10.1 minutes into a stable isotope of carbon with the emission of a positron.

$$_7N^{13*} \rightarrow {}_6C^{13} + {}_1e^0$$

In the case of aluminium,

$$_{13}Al^{27} + {}_2He^4 \rightarrow {}_{15}P^{31*} \rightarrow {}_{15}P^{30*} + {}_0n^1$$

P^{30} is radioactive. The half-life period of radiophosphorus is about 3 minutes and it disintegrates producing a stable isotope of silicon and ejecting a positron.

$$_{15}P^{30*} \rightarrow {}_{14}Si^{30} + {}_1e^0$$

Curie and Joliot verified their explanation by chemical analysis. In the boron reaction, the target used was boron nitride (BN). After irradiation with α-particles for a few minutes, it was heated with caustic soda. One of the products of this chemical reaction was gaseous ammonia NH_3. When the various products were tested for positron activity, it was found that only NH_3 had it. This indicated that nitrogen was the radioelement produced in the experiment. Its half-life was found to be the same as that produced in other irradiated boron targets.

This phenomenon, by which even light elements could be made artificially radioactive, is called *induced radioactivity*. The artificial radioactivity follows the same laws of decay as natural radioactivity. Unlike natural radioactive substances which emit α, β and γ-rays, artificial radioactive substances emit electrons, neutrons, positrons, or γ-rays.

Practically all elements can be made artificially radioactive by bombarding them with various particles. It is found that electron emission takes place in artificial radioelements produced by (n, p), (n, α), (d, p) and (n, γ) reactions.

EXAMPLES

(*i*) $$_7N^{14} + {}_0n^1 \rightarrow ({}_7N^{15}) \rightarrow {}_6C^{14*} + {}_1H^1$$

$$_6C^{14*} \rightarrow {}_7N^{14} + {}_{-1}e^0$$

(*ii*) $$_{13}Al^{27} + {}_0n^1 \rightarrow ({}_{13}Al^{28}) \rightarrow {}_{11}Na^{24*} + {}_2He^4$$

$$_{11}Na^{24*} \rightarrow {}_{12}Mg^{24} + {}_{-1}e^0$$

(*iii*) $$_{15}P^{31} + {}_1H^2 \rightarrow ({}_{16}S^{33}) \rightarrow {}_{15}P^{32*} + {}_1H^1$$

$$_{15}P^{32*} \rightarrow {}_{16}S^{32} + {}_{-1}e^0$$

(*iv*) $$_{49}In^{115} + {}_0n^1 \rightarrow ({}_{49}In^{116}) \rightarrow {}_{49}In^{116*} + \gamma$$

$$_{49}In^{116*} \rightarrow {}_{50}Sn^{116} + {}_{-1}e^0$$

Positron emission takes place in artificial radioelements produced by (p, γ), (p, n), (α, n), (d, n) and (γ, n) reactions.

EXAMPLE

$$_6C^{12} + {}_1H^1 \rightarrow ({}_7N^{13}) \rightarrow {}_7N^{13*} + \gamma$$

$$_7N^{13*} \rightarrow {}_6C^{13} + {}_1e^0 \qquad \text{(T = 13 minutes)}$$

34.10 Preparation of Radioelements

(1) Artificially radioactive elements are now being produced by placing the elements in nuclear reactors. The large number of neutrons available in the nuclear reactor bombard these elements. Thus radiophosphorus can be obtained by neutron bombardment of P^{31}. A large number of isotopes are obtained from nuclear reactors by the neutron radiative capture (n, γ) process.

$$_{15}P^{31} + {}_0n^1 \rightarrow {}_{15}P^{32} + \gamma$$

$$_{27}Co^{59} + {}_0n^1 \rightarrow {}_{27}Co^{60} + \gamma$$

The most important radioisotopes so produced are Tritium, H^3, Radiosodium, Na^{24}, Radiophosphorus, P^{32}, Radio-cobalt, Co^{60}, Radio-carbon, C^{14}, Radio-Iodine, I^{131} and Radio-sulphur, S^{35}.

(2) Another method of preparing a radio-element is to bombard the element with deuterons from a cyclotron. Radio-sodium can be obtained by bombarding Mg^{24} with deuterons. The reaction in this case is (d, α) reaction. The half-life of the radio-isotopes ranges from a fraction of a second to several thousand years.

34.11 Applications of Radioisotopes

Tracer technique. As tracers, radio isotopes have found wide application in agriculture, medicine, industry and research. To study the dynamics of a particular system, a radio-isotope is added and the course of the isotope is studied to understand the whole system. For instance, $_8O^{15}$ mixed with normal air breathed by a patient, enables radio-detection of the path of inhaled oxygen in the lungs. Phosphorus uptake by plants from the soil is studied using a phosphate fertilizer containing the radio-isotope P^{32}. Brain and thyroid tumours are detected using I-131.

(1) **Medical Applications.** Medical applications of radio-isotopes can be divided into two parts: (*i*) *diagnosis* and (*ii*) *therapy*.

Radio-isotopes are used to diagnose the nature of blood circulatory disorders, defects of bone metabolism, to locate tumours etc. For the study of the circulatory disorder in blood vessels, radio-sodium is injected. The constriction in a blood vessel is detected with a counter by surveying the area of the vessel. Let us consider another example. Iodine taken into the body concentrates strongly in the thyroid gland ,and an overactive thyroid holds it much longer than a normal thyroid does. If a patient drinks a glass of water containing I^{131} (as NaI), the count rate in a detector placed near his neck 24 hr later immediately determines whether his thyroid is normal, overactive or underactive. A similar technique is used to diagnose anemia. An anemic patient retains iron in the blood longer than a normal patient. So by injecting a radioactive iron isotope and later withdrawing blood and measuring its activity, anemia can be detected. In therapy, radio-Iodine is used in the treatment of thyroid gland. Radio-phosphorus is used in the treatment of skin diseases. Radio-cobalt and radio-gold are used in the treatment of cancer.

Radio-sodium has been largely used in medical research to study the action of various medicines. A certain amount of $NaCl$ having radioactive sodium in it is mixed with the medicine and administered to the patient. With the help of radiations from radio-sodium, the circulation of the medicine in the body can be detected and hence its effects can be studied.

(2) **Agriculture.** In agriculture, the radio-isotopes help to raise crop yields. For example, in plant physiology P^{32} is incorporated in the fertilizer and added to the soil. The plant and soil are measured from time to time for activity. Phosphorus is taken up by the plant during its growth. It is also possible to find out whether a particular plant requires root feeding or foliage feeding.

(3) In fundamental scientific research, the α, β and γ radiations emanating from radio-isotopes are used for testing of materials. The material under test is interposed between the source

(radioisotope) and a radiation detector. The change in counting rate in the detector gives a measure of the physical properties and structure of the material.

(4) Radio-isotopes are widely used in molecular biology for producing destructive effects. Radiations sterilize pharmaceuticals and surgical instruments. Large doses of radiations kill bacteria and insects.

(5) Certain perishable cereals exposed to radiations remain fresh beyond their normal life span. Very small doses of radiation prevent sprouting and spoilage of onions, potatoes and gram. Storage time can be enhanced by prior irradiation of the food-item.

(6) **Radio-carbon Dating.** One of the interesting properties of radionuclides is that they have constant half-life irrespective of the state of chemical combination of the atoms in which they reside. This property finds applications in *dating* or *age estimation.* ${}_6C^{14}$, a carbon isotope is used in archaeological dating. This isotope has a half-life of 5730 years. The interaction of cosmic rays with the atmospheric gas atoms produces the radioactive isotope C^{14}. This is assimilated by living organism along with the more abundant C^{12}. The production of ${}_6C^{14}$ and its subsequent decay are in equilibrium in the earth's atmosphere. Hence ${}_6C^{14}$ is present in equilibrium concentration in all living plants. However, after the death of the creature or plant, the intake of ${}_6C^{14}$ stops. With the death of the organism, C^{14} begins to decay. Hence the number of C^{14} atoms to the number of stable C^{12} atoms in the object decreases with time after death. Thus the time elapsed since death can then be estimated by measuring this ratio. Using this procedure, ages of wooden slabs from tombs, charcoal used for old paintings etc., have been studied. The same technique has been used in estimating the age of the earth from measurements on the relative amounts of U^{238} and Pb^{206} in geological specimens.

(7) **Industry.** The tracer technique is used in industry, e.g., to know how much a given bearing will wear out under particular conditions of running. The bearing is made of radioactive steel. A measurement of radioactivity in the lubricating oil is a measure of the extent to which wearing out has taken place.

THE NEUTRON

34.12 The Discovery of the Neutron

In 1930, Bothe and Becker found that when beryllium was bombarded with α-particles of energy 5 MeV, a highly penetrating radiation was emitted. This radiation was capable of traversing large thicknessess of lead. It was unaffected by a magnetic field. Hence they assumed that the penetrating radiation consisted of hard γ-rays produced according to the nuclear reaction

$$ {}_4Be^9 + {}_2He^4 \rightarrow {}_6C^{13} + h\nu \qquad ...(1)$$

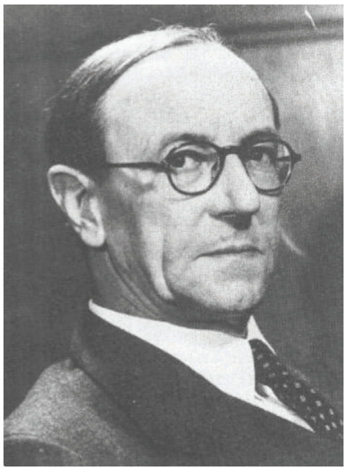
The energy of the γ-rays was calculated by making absorption measurement on it. The energy of the γ-ray was found to be 7 MeV which is greater than the energy of any γ-ray known at that time.

In 1932, Curie and Joliot found that these radiations knocked out protons from paraffin, water and other substances containing hydrogen. The ejected protons had an energy of about 6 MeV. They thought that the protons were ejected as a result of head-on collisions between the γ-ray photons and the protons. They assumed that this photon-proton collision process was analogous to the Compton effect. On this basis, they calculated the photon energy required to oust a 6 MeV proton to be about 60 MeV. This value does not agree with the value 7 MeV obtained from absorption experiments. Further, this value far exceeds the value which results from energy considerations of the nuclear reaction [Equation (1)] mentioned above. Assuming that the α-particle absorbed by beryllium possesses an energy of 5 MeV, the energy $h\nu$ of the resulting γ photon comes out to be about 15 MeV.

To overcome these difficulties, Chadwick suggested that an uncharged particle, called the neutron was emitted in the above reaction. The nuclear reaction is represented by

$$_4Be^9 + _2He^4 \rightarrow _6C^{12} + _0n^1$$

Here $_0n^1$ represents a neutron with zero charge and mass number 1. These neutrons have nearly the same mass as protons and hence can easily knock them out from hydrogenous compounds. The extremely high penetrating power is explained by the fact that they are uncharged and hence experience no repulsion from positive nuclei of lead sheets. Chadwick confirmed his hypothesis by the following experiment.



Sir James Chadwick
(1891 – 1974)

α-particles from polonium deposited on a disc D bombard a beryllium target T (Fig. 34.5). Both the disc and the target are enclosed in an evacuated chamber B. The neutrons released from the reaction escape through the thin wall of this chamber and enter an ionisation chamber C through a thin window W. The ionisation chamber is connected to an amplifier and a counter. Since neutrons are uncharged, neutrons by themselves cannot produce ionisation in the chamber directly. When the radiations from B entered C, very few pulses were registered per minute. The pulses are due to the ions ejected by the neutrons from the walls of the chamber.

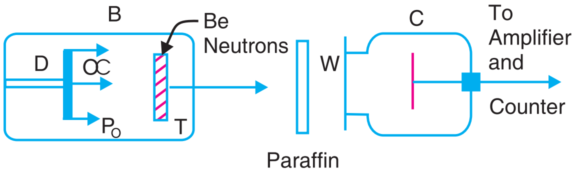


Fig. 34.5

On placing a lead plate in front of W, the number of counts per minute did not increase appreciably. However when a paraffin slab was placed in front of W, the number of counts increased enormously. This is due to the fact that the neutrons eject fast moving protons from paraffin and these protons enter the ionisation chamber through W and are counted.

34.13 Basic Properties of the Neutron

(1) The neutron is one of the fundamental particles constituting the nuclei of atoms.

(2) **Neutron Charge.** Neutrons carry no charge. As they pass through matter, they are neither affected by the positive charge on the nucleus nor by the negative charge on the extranuclear electrons. The neutron, therefore does not lose energy on its passage through matter and is readily able to collide with nuclei bringing about various types of nuclear reactions.

(3) **Decay of the neutron.** In a free state, the neutron is unstable, unlike the proton and the electron. The neutron decays with the emission of a β-particle plus a neutrino, leaving a proton.

$$n \rightarrow p + e^- + \nu + 782\ keV$$

Neutron has a half-life period ($T_{1/2}$) of about 13 minutes. The main difficulty that stands in the way of the experimental determination of the neutron half-life is the fact that the mean-life of free neutrons, before they are captured by nuclei, is short as compared with their half-life.

(4) **Spin and Magnetic moment of the neutron.** Heisenberg suggested that the spin of a neutron may be $\frac{1}{2}h/2\pi$ and this has been confirmed by experiments. We would expect that the magnetic moment of the neutron, which is uncharged should be zero. Actually the magnetic moment of the neutron is about – 1.9 nuclear magneton. According to Yukawa's meson field theory, neutron transforms into a proton and a negative π-meason for a fraction of time. Since this meson is lighter than the proton, its motion will be faster and the resultant magnetic moment is that of a revolving negative charge. Hence it is negative. The proton and π-meson also possess spin. Therefore, the proton and π-meson have intrinsic magnetic moments, that of the meson being large due to its small mass. This situation also contributes to a negative magnetic moment of the neutron.

Assembly for neutron diffraction.

(5) **Neutron diffraction.** According to de-Broglie's theory, all particles are associated with wave characteristics. This wavelength is given by $\lambda = h/mv$ where h is Planck's constant and mv is the momentum of the particle. Thermal neutrons from nuclear reactors have de-Broglie wavelengths of the order 1.82 Å, and are especially suited for crystal structure studies. Neutrons are diffracted by crystals similar to X-rays. Experimental techniques developed for X-ray diffraction can be adopted for this type of work. Wollan has obtained Laue diffraction patterns with thermal neutrons using a sodium choride crystal. Wollan and Stull have studied crystal structure by the powder method using a monochromatic beam of neutrons. The diffracted neutrons were detected by proportional counters filled with BF_3 gas enriched with B^{10}. Neutron diffraction has proved to be a very valuable technique for the study of molecular structure. X-ray and electron diffraction phenomena depend mainly on the number of orbital electrons of the scattering atom. Hence these techniques are not useful to locate the position, in the crystal lattice, of hydrogen atoms, which possess only single orbital electrons.

On the other hand, the diffraction of neutrons results from their scattering by the nucleus and the effect due to this, even with light atoms like hydrogen or beryllium, is quite appreciable. Therefore, neutron diffraction studies make it possible to identify the position of such light atoms in crystals.

Bragg diffraction has been observed in experiments with protons and neutrons. Neutron diffraction is one of the most powerful means of studying crystal structure. Experiments use mono-energetic beams of neutrons and analyse their passage through the crystal. The neutrons emerging from a nuclear reactor through a porthole (Fig. 34.6) have a wide spectrum of energy. The neutron beam is not monochromatic. The neutron beam contains a spectrum composed of many de Broglie wavelengths. When the neutron beam from the reactor falls on the *LiF* crystal, the neutrons observed in the symmetric

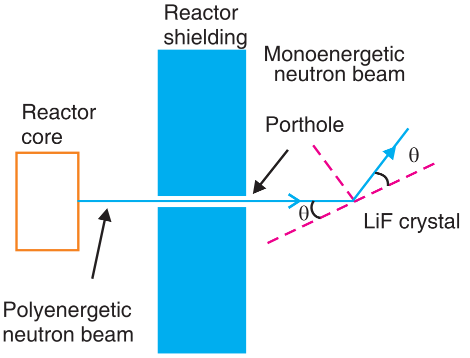


Fig. 34.6

direction correspond only to the wavelength λ given by Bragg's condition $2d \sin\theta = n\lambda$. Therefore they have a well-defined energy and momentum. The crystal then acts as an *energy filter* or *monochromator*. The monoenergetic neutron beam is in turn used to study other materials, by diffraction, or to analyse nuclear reactions involving neutrons of definite energy.

(6) **Neutron mass.** Since the neutron is not a charged particle, its mass could not be determined from deflection measurements in electric and magnetic fields. A precise estimate of the mass of the neutron is obtained from the photo-disintegration of the deuteron discovered by Chadwick and Goldhaber. This method is based on the reaction

$$_1H^2 + \gamma \rightarrow {_1H^1} + {_0n^1}$$

The minimum energy $h\nu$ of the γ photon required to break up the deuteron into a proton and a neutron is known to be 2.225 MeV. The masses of the deuteron and the proton are accurately known from mass spectrograph measurements. Hence m_n can be calculated from the above reaction.

Example. *In the photodisintegration equation for the deuteron the threshold energy is* 2.227 *MeV, viz.*

$$\gamma(h\nu) + {_1H^2} \rightarrow {_1H^1} + {_0n^1} - 2.227\,MeV\ .$$

From the mass spectrometer the mass difference between the doublet $2({_1H^1})$ *and D* $({_1H^2})$*is* 1.5380×10^{-3} *u. If the mass of hydrogen atom is* 1.007825 *u, calculate the mass of the neutron.*

Sol. The mass of the neutron is given by

$$M({_1H^2}) - M({_1H^1}) + \frac{2.227}{931}u = M({_1H^2}) - 2M({_1H^1}) + M({_1H^1}) + 0.002392u$$

$$= -0.001538 + 1.007825 + 0.002392 = 1.008679u.$$

Bell and Elliot used a more accurate method to measure the mass of the neutron. This is based on the reaction

$$_1H^1 + {_0n^1} \rightarrow {_1H^2} + h\nu$$

i.e., when a proton captures a neutron to produce a deutron nucleus, a γ-ray of energy equal to the binding energy is emitted. The γ-ray energy was measured accurately which was found to be 2.230 MeV = 0.002395 amu. We know the masses of $_1H^1$ and $_1H^2$ in mass units and the mass of neutron could be evaluated as follows in atomic mass units

$$_0n^1 = {_1H^2} - {_1H^1} + h\nu = 2.014735 - 1.008142 + 0.002395 = 1.008988 \text{ amu}.$$

In the above measurement, the neutron and proton are initially at thermal energies and no measurement of energy of incident particles is involved.

(7) **Biological effects of neutrons.** It has been found that neutrons produce intense biological effects, even greater than X-rays or γ-rays on normal and tumour tissues. Neutrons are absorbed by fleshy tissues and their destructive action on malignant tissues is very much greater than on normal tissues. It is found that cancerous tissues are destroyed by neutron dosages which are harmless to normal tissues. Hence neutrons are being applied to cancer therapy and patients suffering from cancer are now being treated regularly with neutrons.

(8) **Absorption of neutrons by matter.** The absorption of neutrons by matter is mostly due to their "short range interactions" with nuclei. This interaction may be regarded as a collision which may be either *elastic* or *inelastic*.

(*i*) **Elastic collision.** In an elastic collision, the neutrons are scattered according to the laws of conservation of energy and momentum. A part of the energy of the striking neutron is transferred to the struck nucleus, the lighter the nucleus, the greater being the energy imparted to it. Thus fast neutrons will slow down rapidly when they collide with light nuclei. This explains the slowing down action of hydrogenous materials. Light elements are therefore used to slow down neutrons and these are called *moderators*.

(*ii*) **Inelastic collisions.** This may be subdivided into two categories :

(*a*) *Partial inelastic collisions.* Here the neutron is scattered anomalously and the struck nucleus undergoes internal changes in energy and is raised to an excited state of higher energy. This is very prominent excepting in the case of some very heavy nuclei. This type of scattering may be detected either by measuring the energy of the scattered neutrons or by the γ-rays that are emitted by the nucleus when it returns to a state of lower energy.

(*b*) *Total inelastic collisions.* Here the neutron is captured by the nucleus leading to a real transformation. The capture takes place when the energy of the neutron corresponds to the energy of an excited state of the nucleus. The nucleus and the neutron are now said to be in tune and the capture is called a *resonance capture.* This kind of inelastic collision, known as the *capture process*, plays a prominent role in transmutations produced by neutrons.

34.14 Classification of Neutrons

Neutrons are classified according to their kinetic energy.

(1) **Slow neutrons.** Neutrons with energies from zero to about 1000 eV are usually included in this category. Slow neutrons which are in thermal equilibrium with the medium through which they pass are called *thermal neutrons*. They exhibit a Maxwellian distribution of energy characteristic of the temperature of the medium.

(2) **Fast neutrons.** Neutrons with energies having range between 0.5 to 10 MeV are called fast neutrons. When neutrons pass through a material, they are slowed down by collisions with the nuclei of the material and lose a part of their energy. For light nuclei the loss is large and neutrons are slowed down quickly. Materials rich in hydrogen are very efficient in slowing neutrons and are known as *moderators*. Cadmium has the property of absorbing neutrons and is known as *absorber*.

34.15 Neutron Sources

(*i*) Radioactive sources. Several (α, *n*) reactions using α-particles from natural radioactive substances are used to produce neutrons. The source consists of an intimate mixture of about 5 parts of fine beryllium powder to 1 part of radium. The majority of the neutrons is produced by the (α, *n*) reaction given below :

$$_4Be^9 + {_2He^4} \rightarrow {_6C^{12}} + {_0n^1}$$

This neutron source is isotropic and produces neutrons over a wide energy range (0–13 MeV). Since the half-life of radium is 1600 years, this source is quite steady and gives a long time service.

(*ii*) Photo-neutron Sources. γ-rays from some radioactive substances have been used to produce neutrons by (γ, *n*) reactions in light elements. Neutrons can be produced by the disintegration of beryllium or deuterium by γ-rays by the reactions given below :

$$_4Be^9 + h\nu \rightarrow {_4Be^8} + {_0n^1}$$

$$_1H^2 + h\nu \rightarrow {_1H^1} + {_0n^1}$$

The special feature of these reactions is that they yield mono-energetic neutrons, unlike (α, *n*) reactions from natural α-emitters and produce a flux of 10^7neutrons/sec. The energy of neutrons is, however, low and is not much greater than 1 MeV.

(*iii*) Production of higher energy neutrons by accelerating machines. Neutrons are produced by (*d*, *n*) and (*p*, *n*) reactions in particle accelerators. Van de Graffe machine, cyclotron and other charged particle accelerators serve as efficient neutron generators. Some of the reactions used for production of neutrons are :

$$(d, n)\ \text{reactions}\quad {}_1H^2 + d \rightarrow {}_2He^3 + {}_0n^1 + 3.265\,MeV$$

$${}_1H^3 + d \rightarrow {}_2He^4 + {}_0n^1 + 17.6\,MeV$$

$${}_4Be^9 + d \rightarrow {}_5B^{10} + {}_0n^1 + 3.79\,MeV$$

$$(p, n)\ \text{reactions}\quad {}_1H^3 + p \rightarrow {}_2He^3 + {}_0n^1 - 0.764\,MeV$$

$${}_3Li^7 + p \rightarrow {}_4Be^7 + {}_0n^1 - 1.646\,MeV\ .$$

(*iv*) Reactor sources. The most powerful source of neutrons available in these days is a *nuclear reactor*. The fission of each heavy nucleus by slow neutron capture is accompanied by the release of about 2.5 neutrons on the average. The neutrons produced from the fission of uranium atoms are fast and are slowed down in the moderator. The disadvantages of reactor sources are the wide spread in neutron energy and the impossibility of directly pulsing the neutron source.

34.16 Neutron Detection

Neutrons are uncharged particles and therefore produce negligible ionisation in passing through matter. They cannot be detected by G.M. counters or ionisation chambers or similar other devices which are based on detection by ionisation. Detectors used for neutrons use some indirect process in which neutrons produce charged particles. Also, fission may be utilized for detection of neutrons by placing fissionable material inside an ionization chamber and observing the ionization generated by the fission fragments.

(1) Slow neutron detection—Boron detectors. The neutrons are directed into a chamber filled with boron trifluoride gas. Alternatively, the walls of the chamber are lined with a compound of boron. The boron nucleus absorbs the neutrons and emits α-particles which can be counted. The detection process is based on the reaction,

$${}_5B^{10} + {}_0n^1 \rightarrow {}_3Li^7 + {}_2He^4 + 2.78\,MeV.$$

The cross-section for the above reaction follows the $1/v$ law and falls off with increasing energy. Hence these counters are useful for slow neutrons. Ionisation chambers with boron coated electrodes or proportional counters filled with boron trifluoride vapour (enriched in B^{10}) are in common use.

(2) Fast neutron detection-Proton-recoil detector. This method is based on the observation of ionisation produced by the recoiling protons in the elastic scattering of neutrons by hydrogenous materials. For this purpose, a proportional counter is filled with argon and a thin sheet of a hydrogenous material such as paraffin is placed at one end of the chamber. Fast neutrons striking the target (paraffin) cause the ejection of protons. The protons produce ionisation in their paths through the counter and so can be detected.

34.17 Neutron Collimator

Dunning devised an arrangement known as *neutron howitzer* for collimating neutrons in a beam. He used the property of slowing down of neutrons in hydrogenous material like paraffin. The neutron source *S* (Ra + Be) is placed at the bottom of a barrel inside a paraffin block (Fig. 34.7). A cylindrical slit is cut in this block which is lined from all sides by cadmium sheet. Neutrons emitted in different directions are slowed down by paraffin. When they get slowed down they are effectively absorbed by *Cd* sheet which is very good absorber of slow neutrons. The *Cd* layer is covered by a *Pb* sheet just to absorb γ-rays which are emitted by *Cd* after neutron capture. However, the fast neutrons emitted along the axis of the barrel move outward in a collimated form.

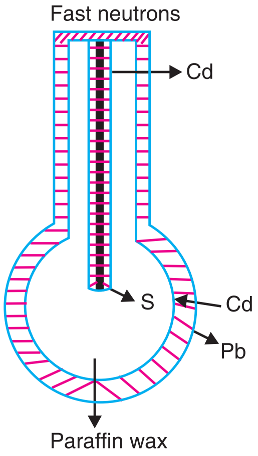


Fig. 34.7

Neutron solar collimator.

EXERCISE

1. Write an essay on the artificial transmutation of elements.
2. What are nuclear reactions? Explain, with typical examples the principles involved in such reactions. What is *Q*-value of a nuclear reaction and explain its significance.
3. What are the advantages of using neutrons as projectiles for effecting artificial transmutations?
4. Write an essay on artificial radioactivity.
5. Give with examples nuclear reactions using different projectiles in which isotopes are produced.
6. Write an essay on the production and application of radio-isotopes.
7. Give one example each of (α, *n*), (*p*, α), (*d*, α), (*d*, *p*), (*n*, α) and (*d*, *n*) reactions. What is photo -disintegration ?
8. A nuclear reaction is given by $_7N^{14} + {_0n^1} \rightarrow {_6C^{14}} + {_1H^1} + 0.55$ MeV. Find the mass of $_6C^{14}$ in u. [**Hint.** Mass of interacting particles = (14.003074 + 1.008665)u = 15.011739u.
   Mass of $_1H^1$ + mass equivalent of 0.55 MeV
   = 1.007825 u + (0.55/931.3) u = 1.0084156 u.
   ∴ mass of $_6C^{14}$ in u = 15.011739u –1.0084156u = 14.0033203u]
9. Write a concise account of the discovery, production, detection, properties and uses of neutrons.
10. Complete the following nuclear reactions :
    (*i*) $_{17}Cl^{35} + ? \rightarrow {_{16}S^{32}} + {_2He^4}$
    (*ii*) $_5B^{10} + ? \rightarrow {_3Li^7} + {_2He^4}$
    (*iii*) $_3Li^6 + ? \rightarrow {_4Be^7} + {_0n^1}$. [**Ans.** $_1H^1$, $_0n^1$, $_1H^2$].
11. When a neutron is absorbed by a target nucleus, the resulting compound nucleus is usually more likely to emit a gamma ray than a proton, deuteron or alpha particle. Why ?
12. How much energy must a photon have if it is to split an α-particle into a triton ($_1H^3$) and a proton? [**Ans.** 18.8 MeV]

13. The rest mass of $_{13}Al^{27}$ is 26.98154u. Find the mass of the product nuclei for the following reactions :

(*a*) Al^{27} (n, γ) Al^{28}, $Q = 7.722$ MeV;

(*b*) Al^{27} (p, α) Mg^{24}, $Q = 1.594$ MeV;

(*c*) Al^{27} (d, p) Al^{28}, $Q = 5.497$ MeV. [**Ans.** 27.98245, 23.98676 and 27.98192 u].

14. Find the threshold energy for the reaction $_7N^{14}$ $(n, \alpha)_5B^{11}$.

[**Hint**. The reaction is $_7N^{14} + {_0n^1} \rightarrow {_5B^{11}} + {_2He^4} + Q$ *i.e.*,

$14.003074 + 1.008665 = 11.009305 + 4.002603 + \text{Q}$ or

$Q = -\,0.000169 \text{ u} = -\,0.16 \text{ MeV}$. $E_{th} = 0.17$ MeV].

15. The cross-section of ^{113}Cd for capturing thermal neutrons is 2×10^4 *b*, the mean atomic mass of natural cadmium is 112 u, and its density is 8.64×10^3 kg/m^3. (*a*) What fraction of an incident beam of thermal neutrons is absorbed by a cadmium sheet 0.1 mm thick ? (*b*) What thickness of cadmium is needed to absorb 99 percent of an incident beam of thermal neutrons ? Cd^{113} isotope constitutes only 12% of natural cadmium. [**Ans.** 0.67, 0.41 mm]

16. The capture cross-section of ^{59}Co for thermal neutrons is 37 *b*.

(*a*) What percentage of a beam of thermal neutrons will penetrate a 1mm sheet of ^{59}Co? The density of ^{59}Co is 8.9×10^3 kg/m^3.

(*b*) What is the mean free path of thermal neutrons in ^{59}Co? [**Ans.** 71.4 per cent ; 2.97 mm]

17. Why do very light atoms reduce the energy of neutrons more than heavy atoms do ?

[**Ans.** If a particle of mass *m* and energy *E* collides with a particle of mass *M*, the maximum amount of energy lost by *m* to *M* in a perfectly elastic collision is $E \times 4\,Mm\,/\,(M + m)^2$. If *M* is much larger than *m*, this is equal to $4\,Em/\,M$ approximately and is a small quantity. But if $M = m$, or nearly so, it equals *E*. This shows that neutrons lose little energy when colliding with large nuclei elastically, but can lose all their energy colliding with protons].

18. The atomic masses of boron 10 and lithium 7 are 10.01294 u and 7.01600 u respectively. The mass of a neutron is 1.00867 u. Show that the release of α-particles from boron 10 can be accomplished by neutrons of thermal energies.

[**Hint:** The nuclear reaction in this process is

$_5B^{10} + {_0n^1} \rightarrow {_3Li^7} + {_2He^4}$

$\Delta\, m = [(10.01294 + 1.00867) - (7.01600 + 4.00260)]\text{u} = 0.00301$ u.

$= 2.802$ MeV.

Thus, in the above reaction 2.802 *MeV* of energy will be released even if the kinetic energy of the neutron is zero. Thermal neutrons have energies of the order of 0.026 *eV* at room temperature (300 *K*). Therefore we conclude that thermal neutrons can initiate the above reaction].

19. Calculate the minimum energy of a photon that will break the nuclide $_3Li^7$ into an α-particle and a triton. The atomic masses are : $_3Li^7 = 7.016004$ u, $_2He^4 = 4.002603$ u, $_1H^3 = 3.016050$ u and $1 \text{ u} \times \text{c}^2 = 931.5$ MeV.

[**Hint.** $_3Li^7 + \gamma \rightarrow {_2He^4} + {_1H^3}$]

The *minimum* energy of the photon must be equal to the rest energy of the product particles minus the (rest) energy of $_3Li^7$.

$$\gamma = [m({_2}He^4) + m({_1}H^3) - m({_3}Li^7)]c^2 = (0.002649\,\text{u})c^2 = \mathbf{2.47\,MeV}.\ ]$$

20. Find the minimum energy required for an α-particle to initiate the following nuclear reaction.

$$_7N^{14} + {_2He^4} + 1.18 \text{ MeV} \rightarrow {_8O^{17}} + {_1H^1}$$

Given : atomic mass of nitrogen -14 = 14.003074 amu

atomic mass of helium - 4 = 4.002604 amu [**Ans.** 1.517 MeV]

21. State the uses of neutrons.

1. The neutrons produce intense biological effects even more than X-rays and γ-rays. That is why, they are used in medicine, specially in the treatment of cancer.

2. Fast and slow neutrons are used for artificial disintegration of nuclei and producing radio-isotopes.
3. Slow neutrons disintegrate Uranium-235 and Pu-239 in nuclear fission.
4. Neutrons are extremely useful for studying the properties of solids. Neutron diffraction studies are superior to those of X-ray diffraction.

22. *Give a brief account of p-p and n-p scattering experiments and discuss their bearing on the nature of nuclear forces.*

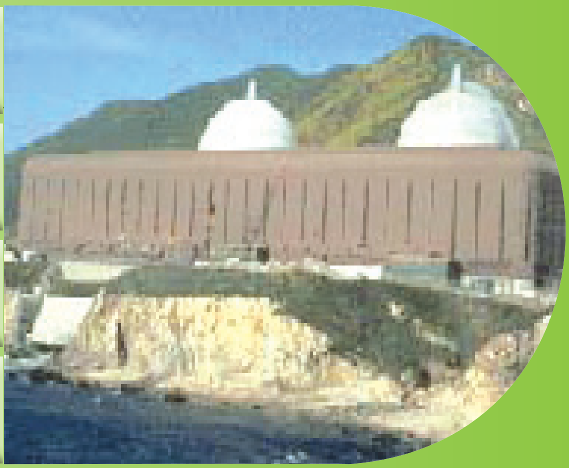
Ans. The most straightforward techniques to study nuclear forces are based on scattering experiments. The study is mainly about the angular distribution of neutrons and protons elastically scattered by protons.

There are two essential differences between the scattering by neutron– proton system and proton-proton system.

(*i*) In *p-p* scattering, in addition to the specific nuclear forces, Coulomb repulsive forces are also present.

(*ii*) Another difference lies in that the scatterer and the scattered particles are *identical*.

The study reveals that the internuclear potential has a hard core that prevents the nucleons from approaching each other much closer than about 0.4 fm. Two nucleons within about 2 fm of each other feel an *attractive nuclear force.* The nuclear force is essentially zero outside to about 3 fm. Since the nuclear force falls to zero very abruptly, it is a *short-range force*. The *p-p* and *p-n* scattering experiments have also revealed the form of the nuclear potential for proton-proton pairs and neutron-proton pairs. The shape of neutron-neutron pairs are determined by indirect methods. It has been found that the nuclear potential for all the pairs are of the same form. This leads to the important conclusion that the nuclear force is independent of the charge of the nucleons. This suggests that the proton and neutron can be considered as different charge states of the same particle.



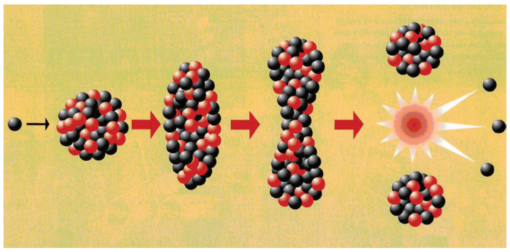
CHAPTER 35

NUCLEAR FISSION AND FUSION

AT A GLANCE			
35.1	Discovery	**35.2**	Nuclear Fission
35.3	Energy Released in Fission	**35.4**	Chain Reaction
35.5	Atom Bomb	**35.6**	Nuclear Reactors
35.7	Nuclear Fusion	**35.8**	Source of Stellar Energy
35.9	Thermonuclear Reactions	**35.10**	Transuranic Elements

35.1 Discovery

In experiments performed by Enrico Fermi, in 1934, the bombardment of uranium by neutrons produced several β-ray activities with different half-lifes. Since uranium decays by α-particle emission with a very long half-life, it was assumed that transuranic elements ($Z > 92$) were being formed. However, in 1938, Hahn and Strassman showed by very careful chemical analysis that one of the radioactive elements produced when uranium is bombarded by neutrons is an isotope of barium, $Z = 56$ ($_{56}Ba^{114}$). They found that barium was accompanied by a radioactive isotope of the gas krypton ($Z = 36$). The atomic numbers of these two nuclides add up to 92, the atomic number of uranium. Hence, Meitner and Frisch suggested that on neutron bombardment, the uranium nucleus spilts up into two lighter nuclei (Ba and Kr). The phenomenon was called *fission*, since it resembled division of cells in biology.



Nuclear Fission.

35.2 Nuclear Fission

The process of breaking up of the nucleus of a heavy atom into two, more or less equal fragments with the release of a large amount of energy is known as fission.

When uranium is bombarded with neutrons, a uranium nucleus captures a slow neutron,forming an unstable compound nucleus. The compound nucleus splits into two nearly equal parts. Some neutrons are also released in this process.

The schematic equation for the fission process is

$$_{92}U^{235} + {}_0n^1 \rightarrow {}_{92}U^{236*} \rightarrow X + Y + neutrons \qquad ...(1)$$

$_{92}U^{236*}$ is a highly unstable isotope, and X and Y are the fission fragments. The fragment are not uniquely determined, because there are various combinations of fragments possible and a number of neutrons are given off. Typical fission reactions are

$$_{92}U^{235} + {}_0n^1 \rightarrow {}_{92}U^{236*} \rightarrow {}_{56}Ba^{141} + {}_{36}Kr^{92} + 3\,{}_0n^1 + Q \qquad ...(2)$$

$$_{92}U^{235} + {}_0n^1 \rightarrow {}_{92}U^{236*} \rightarrow {}_{54}Xe^{140} + {}_{38}Sr^{94} + 2\,{}_0n^1 + Q \qquad ...(3)$$

where Q is the energy released in the reaction.

According to Eqn. (2), when $_{92}U^{235}$ is bombarded by a slow moving neutron, the nucleus becomes unstable ($_{92}U^{236*}$) and splits into $_{56}Ba^{141}$ and $_{36}Kr^{92}$ releasing 3 neutrons and energy Q (Fig. 35.1).

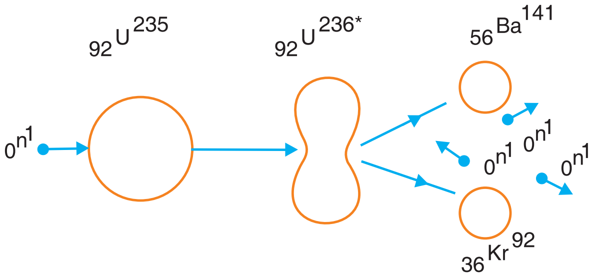


Fig. 35.1

35.3 Energy Released in Fission

We know that in the process of nuclear fission a large amount of energy is released. This energy is produced because the original mass of the nucleus is greater than the sum of the masses of the products produced after fission. The difference between these masses before and after fission is converted into energy according to Einstein' s equation $E = mc^2$.

The energy liberated per fission can be calculated as follows :

Let us consider the fission of $_{92}U^{235}$. The fission reaction is shown in Fig. 35.1.

(*i*) From the binding energy curve (Fig. 27.3), for a heavy nucleus (with A ~ 236), the average binding energy per nucleon $\overline{B}$ is ~ 7.6 MeV. For a nucleus with a medium weight, it is about 8.5 MeV. Hence, the average binding energy per particle is about 0.9 MeV more in the fission products than in a compound nucleus (*e.g.*, $^{236}_{92}\mathrm{U}$) . Thus, this excess of 0.9 MeV per nucleon must be liberated in the fission process. The total energy liberated per fission of $^{236}_{92}\mathrm{U}$ is given by

$$E_f = 236\ (8.5 - 7.6)\ \mathrm{MeV} \sim 212\ \mathrm{MeV}.$$

(*ii*) We shall calculate the total energy released by considering only the fission products (i.e., stable end products). Let us consider the reaction

$$n + {}^{235}_{92}\mathrm{U} \rightarrow {}^{98}_{42}\mathrm{Mo} + {}^{136}_{54}\mathrm{Xe} + 2n\ .$$

The total mass of the nuclei before fission is

1.0087 + 235.0439 = 236.0526 u.

The total mass of the nuclei after fission is

97.906 + 135.9072 + 2.0174 = 235.8306 u.

The mass difference is 0.2220 u.

The energy release = 0.2220 × 931.3 = 206.7 MeV.

Energy released by 1kg of uranium.

Number of atoms in 1kg of uranium = $(6.023 \times 10^{26}) / 235$

Energy released in one fission = 200 MeV

$$\left.\begin{array}{l}\text{Energy produced}\\ \text{by 1 kg of uranium}\\ \text{during fission}\end{array}\right\} = E = \frac{6.023\times10^{26}}{235}\times200 = 5.128\times10^{26}\ \text{MeV}$$

$$\therefore\ E = (5.128\times10^{26})\times(1.6\times10^{-13})\ \text{J}\ (\because\ 1\ \text{MeV} = 1.6\times10^{-13}\ \text{J})$$

$$= \frac{(5.128\times10^{26})\times(1.6\times10^{-13})}{3.6\times10^{6}}\ \text{k Wh}\ \ (\because\ 1\ \text{k Wh} = 3.6\times10^{6}\ \text{J})$$

$$= 2.26\times10^{7}\ \text{k Wh.}$$

Thus the energy released by fission of 1 kg of U^{235} is 2.26×10^7 kWh. Due to this reason, nuclear energy is being used for the generation of electricity. The most striking aspect of nuclear fission is the magnitude of the energy involved. Ordinary chemical reactions, such as those that participate in the combustion of coal and oil, liberate only a few electron volts per individual reaction. Most of the energy that is released during fission goes into the K.E. of the fission fragments.The emitted neutrons, β and γ -rays and neutrinos carry off perhaps 20% of the total energy.

Bohr and Wheeler's theory of nuclear fission. Bohr and Wheeler successfully explained the phenomenon of nuclear fission on *liquid drop model*. A liquid drop has a spherical form due to internal molecular forces resposible for surface tension. According to their theory, an excited liquid drop may oscillate in a number of ways. On applying a large external force, the sphere may change into an ellipsoid. If the external force is sufficiently large, the ellipsoid may change into a dumb-bell shape and may even break at the narrow end into two portions.

The analogy may be extended to a nucleus which behaves like a liquid drop. When a nucleus absorbs a neutron, it forms a compound nucleus which is highly energetic (Fig. 35.2). The extra energy possessed by it comes mostly from the binding energy of the neutron absorbed by it. The extra energy may set up a series of rapid oscillations in the spherical compound nucleus shown as *A* in Fig. 35.2. As a result of these oscillations, the shape of the nucleus may change at times from spherical to ellipsoidal shown as *B*. If the extra energy is large, oscillations may be so violent that stage *C* and ultimately stage *D* may be approached. The nucleus is now dumbbell shaped (stage *D*). Each bell of the dumb-bell has now a positive charge and one repels the other. This results in a fission (stage *E*).

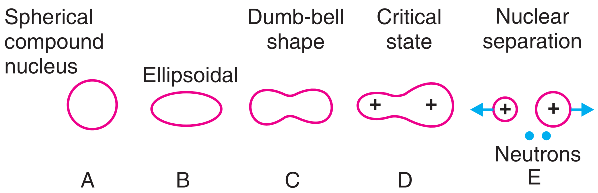


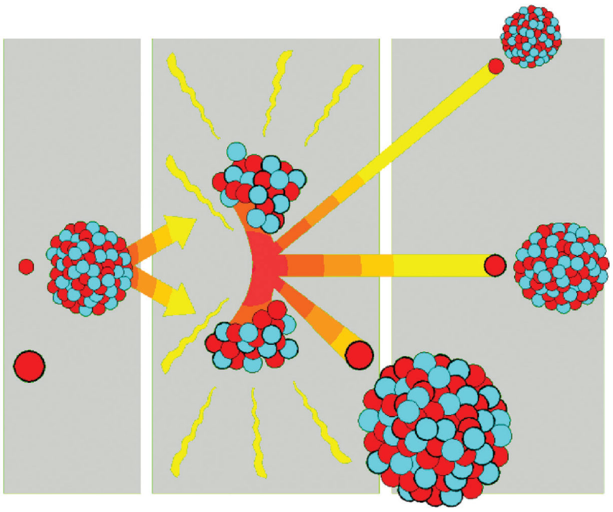
Fig. 35.2

The nuclei that result from fission are called *fission fragments*. Usually fission fragments are of unequal size. A heavy nucleus undergoes fission when it acquires enough excitation energy to oscillate violently. Certain nuclei, notably ${}_{92}U^{235}$are sufficiently excited by the mere absorption of an individual neutron. Other nuclei, notably ${}_{92}U^{238}$ require more excitation energy for fission than the binding energy released when another neutron is absorbed. ${}_{92}U^{238}$ undergoes fission only by reaction with fast neutrons whose kinetic energies exceed about 1 MeV.

35.4 Chain Reaction

A chain reaction is a self-propagating process in which number of neutrons goes on multiplying rapidly almost in geometrical progression during fission till whole of fissile material is disintegrated.

Example : Suppose a single neutron causing fission in a uranium nucleus produces 3 prompt neutrons. The three neutrons in turn may cause fission in three uranium nuclei producing 9 neutrons. These nine neutrons in turn may cause fission in nine uranium nuclei producing 27 neutrons and so on. The number of neutrons produced in n such generations is 3^n. The ratio of secondary neutrons produced to the original neutrons is called the *multiplication factor* (k).



Chain Reaction.

Consider 1 kg of ${}_{92}U^{235}$which contains $6.023 \times 10^{26}/\ 235$ or about 25×10^{23} atoms. Suppose a stray neutron causes fission in a uranium nucleus. Each fission will release on the average 2.5 neutrons. The velocity of a neutron among the uranium atoms is such that a fission capture of a thermal neutron by the ${}_{92}U^{235}$nuclei takes place in about 10^{-8} s. Each of these fissions, in turn, will release 2.5 neutrons. Let us assume that all these neutrons are available for inducing further fission reactions. Let n be the number of stages of fission captures required to disrupt the entire mass of 1 kg of ${}_{92}U^{235}$. Then

$$(2.5)^n = 25 \times 10^{23} \text{ or } n \approx 60.$$

The time required for 60 fissions to take place = $60 \times 10^{-8}\ s = 0.6\ \mu\ s$.

Since each fission releases about 200 MeV of energy, this means that a total of $200 \times 25 \times 10^{23} = 5 \times 10^{26}$ MeV of energy is released in 0.6 μ s. The release of this tremendous amount of energy in such a short time interval leads to a violent explosion. This results in powerful air blasts and high temperature of the order of 10^7 K or more, besides intense radioactivity. The self-propagating process described here is called a *chain reaction.*

Two types of chain reaction are possible. In one, the chain reaction is first accelerated so that the neutrons are built up to a certain level and there after the number of fission producing neutrons is kept constant. This is *controlled chain reaction*. Such a controlled chain reaction is used in nuclear reactors. In the other type of chain reaction, the number of neutrons is allowed to multiply indefinitely and the entire energy is released all at once. This type of reaction takes place in atom bombs.

Multiplication factor (k). The ratio of secondary neutrons produced to the original neutrons is called the multiplication factor. It is defined as

$$k = \frac{\text{Number of neutrons in any one generation}}{\text{Number of neutrons in the preceding generation}}.$$

The fission chain reaction will be "*critical*" or steady when $k = 1$, it will be building up or "*Supercritical*" when $k > 1$ and it will be dying down or "*subcritical*" when $k < 1$.

Critical size for maintenance of chain reaction. Consider a system consisting of uranium (as fissile material) and a moderator. Even though each neutron that produces fission ejects 2.5 neutrons on an average, all of them are not available for further fission. The maintenance of the chain reaction depends upon a favourable balance of neutrons among the three processes given below :

(1) The fission of uranium nuclei which produces more neutrons than the number of neutrons used for inducing fission.

(2) Non-fission processes, including the radiative capture of neutrons by the uranium and the parasitic capture by the different substances in the system and by impurities.

(3) Escape or leakage of neutrons through the surface of the system.

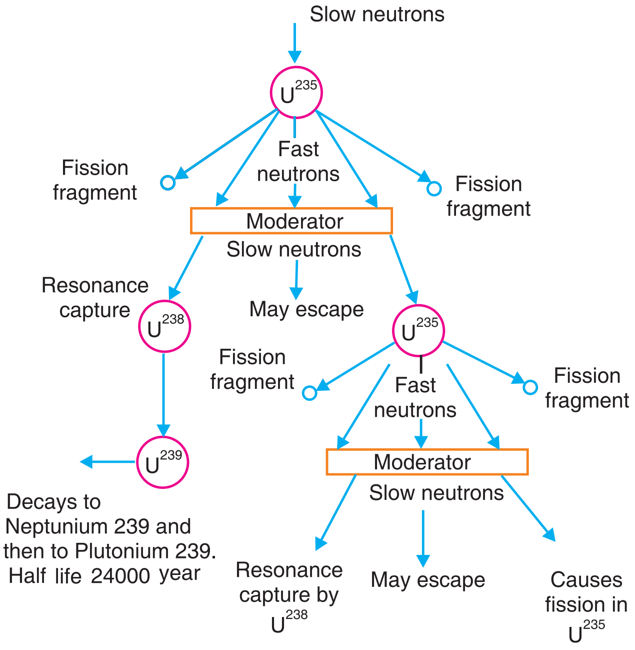
If the loss of neutrons due to the last two causes is less than the surplus of neutrons produced in the first, a chain reaction takes places. Otherwise it cannot take place.

The escape of neutrons takes place from the surface of the reacting body and fission occurs throughout its *volume.*

$\therefore$ Escape rate varies as r^2 and production rate varies as r^3.

$$\therefore \qquad \frac{\text{Escape rate}}{\text{Production rate}} \propto \frac{1}{r}.$$

The larger the size of the body, the smaller is the escape rate. Thus it is clear that by increasing the volume of the system, the loss of neutrons by escape from the system is reduced. The greater the size of the system, the lesser is the probability of the escape of neutrons. In this case, the production of neutrons will be more than the loss due to other causes and a chain reaction can be maintained. Thus there is a critical size for the system. *Critical size of a system containing fissile material is defined as the minimum size for which the number of neutrons produced in the fission process just balance those lost by leakage and non-fission capture.* The mass of the fissionable material at this size is called the *critical mass*. If the size is less than the critical size, a chain reaction is not possible.


Slow neutrons
U235
Fission fragment
Fast neutrons
Fission fragment
Moderator
Slow neutrons
Resonance capture
U238
May escape
U235
Fission fragment
Fast neutrons
Fission fragment
U239
Moderator
Slow neutrons
Decays to Neptunium 239 and then to Plutonium 239. Half life 24000 year
Resonance capture by U238
May escape
Causes fission in U235

(Self-sustaining chain reaction)

Fig. 35.3

Natural uranium and chain reaction. Natural uranium consists of 99.28% of U^{238} and 0.72% of U^{235}. As most of the mass of natural uranium consists of U^{238}, the neutrons released during nuclear fission will try to bombard the nuclei of U^{238} mostly and very few will bombard U^{235}. U^{235} undergoes fission even by neutrons of small energy like thermal neutrons. U^{238} is fissionable only with fast neutrons of energy 1 MeV or more. It has been found that very few neutrons can cause fission of U^{238} but neutrons of all possible energies can cause fission of U^{235}. Thus chain reaction is not possible in natural uranium.

A chain reaction can, however, be made to develop in natural uranium, if the fast neutrons from it are quickly reduced to thermal ones before they are lost through non-fission capture in the uranium, so that the chances of the thermal neutron fission of U^{235} go up. The neutron can be slowed down by distributing among lumps or rods of uranium a meterial called *moderator.* The moderators must not absorb the neutrons. The function of the moderator is to slow down the neutrons produced by fission by elastic collision. Materials used as neutron moderators have a large inelastic scattering cross-section and, at the same time, a small neutron-capture cross-section. Commonly used moderators are graphite, heavy water (D_2O), beryllium, beryllium oxide, hydrides of metals and organic liquids. The nuclei of these substances absorb neutrons only to a slight extent.

Fig. 35.3 shows a self-sustaining chain reaction. A slow neutron bombards a U^{235} nucleus. The nucleus breaks into two fragments and in the fission process three fast neutrons are emitted. The neutrons are slowed down by the moderator. One neutron may escape. One neutron may be captured by U^{238} to form U^{239} which decays to Np^{239}, and then to Pu^{239}. One neutron is still available for carrying on the chain reaction. It bombards U^{235} and the process is repeated.

35.5 Atom Bomb

The principle of fission is made use of in the construction of the atom bomb. An atom bomb consists essentially of two pieces of ${}_{92}U^{235}$ (or ${}_{94}Pu^{239}$) each smaller than the critical size and a source of neutrons. The two subcritical masses of U^{235} in the form of hemispheres are kept apart by using a separator aperture (Fig. 35.4). When the bomb has to be exploded, a third well fitting cylinder of U^{235} (whose mass is also less than the critical mass) is propelled so that it will fit in or fuse together with the other two pieces. Now the total quantity of U^{235} is greater than the critical mass. Hence an uncontrolled chain reaction takes place resulting in a terrific explosion.

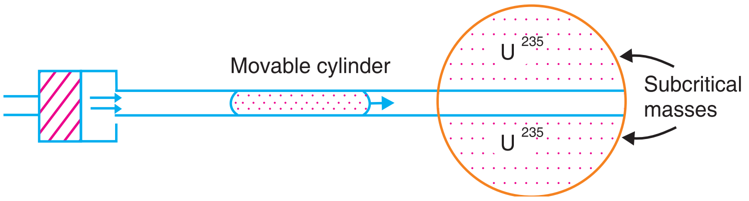


Fig. 35.4

The explosion of an atom bomb releases tremendously large quantity of energy in the form of heat, light and radiation. A temperature of millions of degrees and a pressure of millions of atmospheres are produced. Such explosions produce shock waves. They are very dangerous because the waves spread radioactivity in air and cause loss of life. The release of dangerously radioactive γ-rays, neutrons and radioactive materials presents a health hazard over the surroundings for a long time. The radioactive fragments and isotopes formed out of explosion adhere to dust particles thrown into space and fall back to earth causing a radiation "*fall-out*", even at very distant places.

Atom Bomb.

35.6 Nuclear Reactors

Nuclear Reactor.

We know that during the fission of U^{235} a large amount of energy is released. The atom bomb is due to an uncontrolled chain reaction. A very large amount of energy is liberated within an extremely small interval of time. Hence it is not possible to direct this energy for any useful purpose. But in a nuclear reactor, the chain reaction is brought about under controlled conditions. If the chain reaction is put under control, after some time a steady state is established. Under a steady state, the rate of energy production also attains a constant level. Such a device in which energy is released at a given rate is known as a *nuclear reactor*.

Nuclear reactors consist of five main elements:

(1) The fissionable material called fuel,

(2) Moderator,

(3) Neutron reflector,

(4) Cooling system, and

(5) The safety and control systems.

(1) The fissionable substance. The commonly used fissionable materials are the uranium isotopes U^{233}, U^{238}, the thorium isotope Th^{232}, and the plutonium isotopes Pu^{239}, Pu^{240} and Pu^{241}.

(2) Moderator. The function of the moderator is to slow down the highly energetic neutrons produced in the process of fission of U^{235} to thermal energies. Heavy water (D_2O), graphite, beryllium, etc., are used as moderators. Ideally, moderators have low atomic weight and low absorption cross-section for neutrons.

(3) Neutron reflector. By the use of reflectors on the surface of reactors, leakage of neutrons can be very much reduced and the neutron flux in the interior can be increased. Materials of high scattering cross-section and low absorption cross-section are good reflectors.

(4) Cooling system. The cooling system removes the heat evolved in the reactor core. This heat is evolved from the K.E., of the fission fragments when they are slowed down in the fissionable substance and moderator. The coolant or heat-transfer agent (water, steam, He, CO_2, air and certain molten metals and alloys) is pumped through the reactor core. Then, through a heat exchanger, the coolant transfers heat to the secondary thermal system of the reactor.

(5) Control and safety system. The control systems enable the chain reaction to be controlled and prevent it from spontaneously running away. This is accomplished by pushing control rods into the reactor core. These rods are of a material (boron or cadmium) having a large neutron-absorption cross-section. These rods absorb the neutrons and hence cut down the reactivity. By pushing in the rods, the operation of the reactor can be made to *die down*, by pulling them out to *build up*. The safety systems protect the space surrounding the reactor against intensive neutron flux and gamma rays existing in the reactor core. This is achieved by surrounding the reactor with massive walls of concrete and lead which would absorb neutrons and gamma rays.

Power reactor. The heat generated in a nuclear reactor is used for producing power in a nuclear power plant. Fig. 35.5 shows the components of a power reactor. A quantity of encriched uranium in the form of pure metal or solution of a soluble salt in water constitutes the centre of the heat energy source. A large quantity of heat is produced in the fission process. The cadmium rods regulate the temperature, to a pre-determined value. If it is desired to bring down the temperature, the cadmium rods are pushed down further as to absorb more neutrons. If the temperature has to be raised, the cadmium rods are pulled up a little. A fluid is circulated through the shielded reactor and heat exchanger. The hot fluid, while flowing through the heat exchanger, converts water into steam. The steam produced runs conventional turbines to produce electricity.

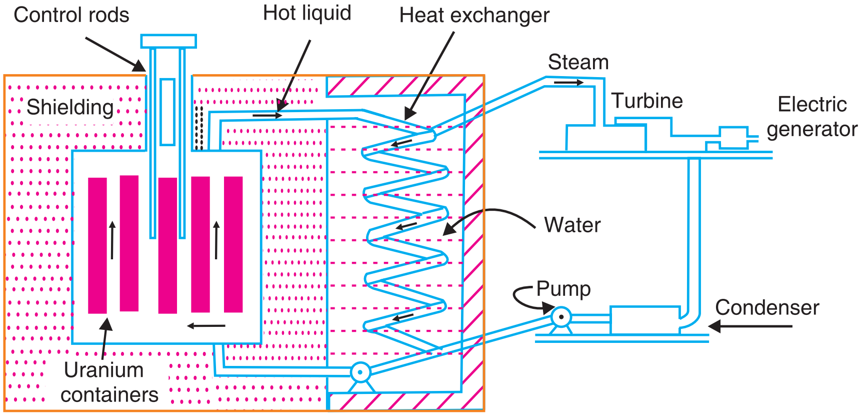


Fig. 35.5

Breeder Reactor. If a thermal reactor core with U^{235} fuel is surrounded by a blanket of a fertile material like U^{238}, U^{238} can be converted into fissile fuel. Reactors of this type are called fuel producing reactors. The reactions are as follows :

$${}_{92}U^{238} + {}_{0}n^{1} \to {}_{92}U^{239} + \gamma$$

This is followed by $\quad {}_{92}U^{239} \to {}_{93}Np^{239} + \beta^-$

${}_{93}Np^{239}$ is also radioactive. It emits a β-particle to form plutonium.

$${}_{93}Np^{239} \to {}_{94}Pu^{239} + \beta^- + \gamma\,.$$

This process of producing one type of fissionable material (Pu^{239}) from a non-fissionable material (U^{238}) is called *breeding* and the reactor a *breeder reactor.*

The breeding reactions for the fertile material ${}^{232}_{90}Th$ are

$${}^{232}_{90}Th + n \to {}^{233}_{90}Th^* \xrightarrow[22\,\text{min}]{\beta^-} {}^{233}_{91}Pa \xrightarrow[27\text{d}]{\beta^-} {}^{233}_{92}U\,(T_{1/2} = 1.6\times10^5 \text{ yrs})\,.$$

Uses of Nuclear Reactors.

(1) Nuclear power. Nuclear reactors are used in the production of electric energy.

(2) Production of radioisotopes. Nuclear reactors are useful in producing a large number of radio-isotopes. To produce radio-isotope, a suitable compound is drawn into the centre of the reactor core where the flux of neutrons may well be more than 10^{16} /m^2/ sec. Sodium-24 is manufactured in this way.

$$_{11}Na^{23} + {}_0n^1 \rightarrow {}_{11}Na^{24*}$$

(3) Scientific research. Reactors produce a number of radioactive materials needed for research purposes. The reactors provide a huge source of neutrons. Using these neutrons, several useful radioisotopes have been artificially produced and several nuclear reactions have been studied. We may also study the effect of neutrons on biological tissues. Reactors may also be used to study radiation damage.

35.7 Nuclear Fusion

Nuclear fusion. In this process, two or more light nuclei combine together to form a single heavy nucleus. For example, when four hydrogen nuclei are fused together, a helium nucleus is formed. The mass of the single nuleus formed is always less than the sum of the masses of the individual light nuclei. The difference in mass is converted into energy according to Einstein's equation $E = mc^2$.

EXAMPLE. Consider a single helium nucleus formed by the fusion of two deuterium nuclei. Mass of $_1H^2 = 2.014102u$; mass of $_2He^4 = 4.002604u$.

$$_1H^2 + {}_1H^2 \rightarrow {}_2He^4$$

The initial mass of 2 deuterium atoms = 2 × 2.014102 = 4.028204 u.

Mass of helium atom = 4.002604 u.

Decrease in mass = 4.028204 – 4.002604 = 0.025600 u.

∴ Energy released = 0.025600 × 931.3 MeV = 23.84 MeV.

Thus the energy released in fusion is 23.84 MeV.

35.8 Source of Stellar Energy

The temperature of the stars are very high and they radiate tremendous amount of energy. The sun is one of the innumerable stars. The sun radiates 3.8×10^{26} joules of energy each second. The origin of such a tremendous amount of energy is neither chemical nor gravitational. The fusion of protons is supposed to release the energy in the sun and in other stars. Bethe suggested the following carbon-nitrogen cycle as one of the most important nuclear reactions for release of energy by fusion.

Carbon-Nitrogen Cycle. The cycle is as follows (Fig. 35.6).

$$_6C^{12} + {}_1H^1 \rightarrow {}_7N^{13*} + \gamma \qquad ...(1)$$

$$_7N^{13*} \rightarrow {}_6C^{13} + {}_1e^0 + \nu \qquad ...(2)$$

$$_6C^{13} + {}_1H^1 \rightarrow {}_7N^{14} + \gamma \qquad ...(3)$$

$$_7N^{14} + {}_1H^1 \rightarrow {}_8O^{15*} + \gamma \qquad ...(4)$$

$$_8O^{15*} \rightarrow {}_7N^{15} + {}_1e^0 + \nu \qquad ...(5)$$

$$_7N^{15} + {}_1H^1 \rightarrow {}_6C^{12} + {}_2He^4 \qquad ...(6)$$

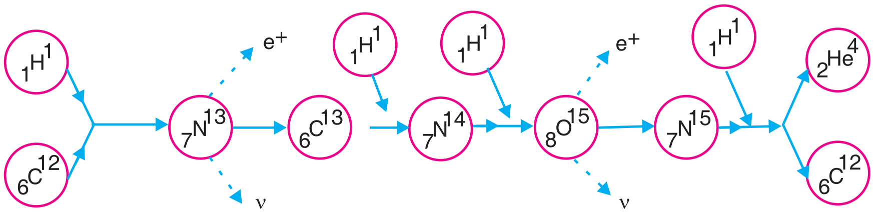


Fig. 35.6

In this cycle C^{12} acts like a catalyst.

The reaction cycle is essentially the reaction

$$4\,{}_1H^1 \rightarrow {}_2He^4 + 2\,{}_1e^0 + 2\nu + Q$$

The loss in mass is calculated as follows :

$$4\ {}_1H^1 = 4.031300;\ {}_2He^4 = 4.002603 \text{ and } 2\ {}_1e^0 = 0.001098.$$

$$\therefore \quad \text{Loss in mass} = 0.02756\ u.$$

$$\therefore \quad \text{Energy released} = 0.02756 \times 931 = 27.5 \text{ MeV}.$$

It is found that in one million years the sun loses about 10^{-7} of its mass by the above process. Taking mass of the sun as 2×10^{30} kg and its present age as 10^{10} years, it is estimated that the *C–N* cycle may keep going for another 30 billion years.

Proton-Proton Cycle. Recent modification of the estimates of the central temperature of the sun now favour the proton-proton chain. In the *p–p* chain, two protons first fuse to produce a deuterium nucleus which combines with another proton to yield He^3. Two He^3 nuclei interact and form He^4 and two protons. These reactions can be represented by the equations,

$$\begin{array}{c} {}_1H^1 + {}_1H^1 \rightarrow ({}_2He^2) \rightarrow {}_1H^2 + {}_1e^0 + \nu + 0.42\,\text{MeV} \\ {}_1H^2 + {}_1H^1 \rightarrow ({}_2He^3) \rightarrow {}_2He^3 + \gamma + 5.5\,\text{MeV} \\ \underline{{}_2He^3 + 2He^3 \rightarrow ({}_4Be^6) \rightarrow {}_2He^4 + {}_1H^1 + {}_1H^1 + 12.8\,\text{MeV}} \\ 4\,{}_1H^1 \rightarrow {}_2He^4 + 2\ {}_{+1}e^0 + 2\nu + 2\gamma + 26.7\,\text{MeV} \end{array}$$

35.9 Thermonuclear Reactions

The source of stellar energy is fusion. This suggests that a large amount of energy can be obtained by nuclear fusion. But it is not easy to fuse the light nuclei into a single nucleus. The main difficulty in the fusion of nuclei is the electric force of repulsion between the positively charged nuclei. Fusion is possible when the K.E. of each of the nuclei is large enough to overcome the repulsion. Fusion reactions can take place only at very high temperatures (of the order of 10^7 to 10^9 K). Only at these very high temperatures, the nuclei are able to overcome their mutual Coulomb repulsion and enter the zone of nuclear attractive forces. Hence these reactions are called *thermonuclear reactions.*

A star is able to control thermonuclear fusion in its core because of its strong self-gravity. The thermonuclear reactions in the core of the sun cause high temperatures which generate strong outward pressures; these act against the sun's own gravity, preventing it from contracting, and holding it in equilibrium. The equation of stellar structure, set up by A.S. Eddington, relates the gravitational force in the star to the progressive changes of pressure from its centre outwards, the magnitude of pressure to density and temperature, and the fall of temperature outwards to the flow of energy from the interior to the surface. From these equations, stable models of stars emerge, with central temperatures high enough to start and sustain thermonuclear fusion. The key role, of course, is played by the controlling force of gravity. The large mass of an astronomical system makes gravity the most important factor in determining its behaviour.

Hydrogen Bomb.

Hydrogen bomb. Hydrogen bomb is a device which makes use of the principle of nuclear fusion. The very high temperature required for an uncontrolled thermonuclear reaction is obtained by the detonation of an atom bomb. In this weapon, hydrogen is the core. The fission bomb produces a very high temperature, at which thermonuclear reactions start resulting in the fusion of hydrogen nuclei to form helium. Greater energy per unit mass is obtained from a hydrogen bomb than from a nuclear fission bomb.

Controlled thermonuclear reactions. A large amount of energy is released in a fraction of a second in a hydrogen bomb. If the thermonuclear reaction could be controlled to take place more slowly, the energy released can be used for constructive purposes. We know that very high temperatures are needed to bring about a nuclear fusion process. The main problem is to produce such a high temperature and to find a container for the gas which can stand this temperature. At this temperature the gas is highly ionised and is called *plasma.* One of the severe engineering problems is the design of a "*container*" in which a very hot plasma can be contained under high pressure to initiate a fusion reaction. Since almost any container would melt in the presence of a plasma, attempts are being made to contain and control plasmas trapped in a specially shaped magnetic field (Fig. 35.7). By increasing the field and changing the shape of the field, it is hoped that the plasma in this "*magnetic bottle*" can be raised to the required temperature and pressure for fusion reactions.

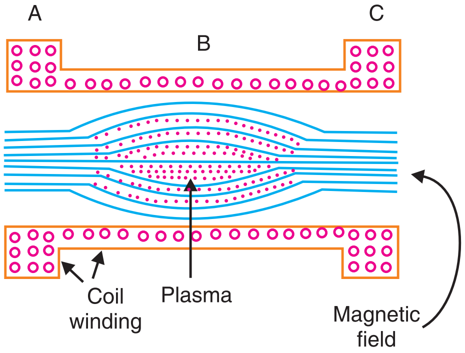


Fig. 35.7

Nuclear fusion as an energy source will be a boon to humanity because of the following reasons:

(1) Hydrogen is available everywhere on this planet in various forms.

(2) The lightness of the reactant nuclei makes the energy yield per unit mass of the reacting material much greater than that in nuclear fission process.

(3) A fusion reactor does not leave behind as in fission reactor radioactive waste, the disposal of which poses a tremendous problem.

EXAMPLE 1. *A reactor is developing energy at the rate of* 3000 *kW. How many atoms of* U^{235} *undergo fission per second ? How many kilograms of* U^{235} *would by used in* 1000 *hours of operation assuming that on an average energy of* 200 *MeV is released per fission ?*

SOL. Rate of development of energy by the reactor = 3000 kW = 3×10^6 J s^{-1}

Energy released per fission = 200 MeV = $200 \times 1.6 \times 10^{-13}$ J

$= 32 \times 10^{-12}$ J.

$$\therefore \text{ No. of atoms undergoing fission per second} = \frac{3 \times 10^6}{32 \times 10^{-12}} = 9.4 \times 10^{16}$$

No. of atoms undergoing fission in 1000 hours = $(9.4 \times 10^{16}) \times (1000 \times 60 \times 60)$

$= 3.384 \times 10^{23}$

According to Avogadro's hypothesis, 6.025×10^{26} atoms of U^{235} weight 235 kg.

$$\therefore \text{ Weight of } 3.384 \times 10^{23} \text{ atoms of } U^{235} = \frac{235 \times (3.384 \times 10^{23})}{6.025 \times 10^{26}}$$

$= 0.1321$ kg.

EXAMPLE 2. *A deuterium reaction that occurs in experimental fusion reactor is* $H^2(d, p)\, H^3$ followed by $H^3(d, n)\, He^4$. (*a*) *Compute the energy release in each of these.* (*b*) *Compute the total energy release per gram of the deuteron used in the fusion.* (*c*) *Compute the percentage of the rest mass of deuteron released as energy.* (*d*) *Compare* U^{235} *fission with deuteron fusion as a source of energy release. Given*

$H^2 = 2.014102\ u$, $H^3 = 3.016049\ u$, $H^1 = 1.007825\ u$, ${}_0n^1 = 1.008665\ u$,

$He^4 = 4.002604\ u$ *and* $U^{235} = 235.0439\ u$.

SOL. (*a*) (*i*) The fusion reaction $H^2(d, p)\, H^3$ is

$$_1H^2 + {}_1H^2 \rightarrow {}_1H^3 + {}_1H^1 + Q$$

Mass decrease in the reaction

$\Delta m = (2.014102 + 2.014102 - 3.016049 - 1.007825)\ u = 0.00433u.$

$\therefore$ Energy released = 0.00433×931.3 MeV = 4.032 MeV.

(*a*) (*ii*) The reaction $H^3(d, n)\, He^4$ is

$$_1H^3 + {}_1H^2 \rightarrow {}_2He^4 + {}_0n^1 + Q$$

$\Delta m = (3.016049 + 2.014102 - 4.002604 - 1.008665)\ u = 0.018842$ u.

Energy released = 0.018842×931.3 = 17.58 MeV.

$\therefore$ Total energy release = 4.032 + 17.58 = 21.61 MeV.

(*b*) This total energy release is from the fusion of 3 ${}_1H^2$ nuclei.

$\therefore$ Energy release per H^2 = 21.61/3 = 7.203 MeV.

$$\text{No. of nuclei in 1 gram of } H^2 = \frac{6.02 \times 10^{23}}{2.014102}$$

$$\therefore \text{ Total release of energy from 1 gram of } H^2 = \frac{6.02 \times 10^{23}}{2.014102} \times 7.023 = 2.153 \times 10^{24} \text{ MeV}.$$

(*c*) Energy equivalent of one H^2 nucleus = 2.014102 × 931.3 MeV

Average release of energy per H^2 nucleus = 7.023 MeV.

$$\therefore \left.\begin{array}{l}\text{The percentage of the rest of mass of}\\ \text{deuteron released as energy}\end{array}\right\} = \frac{7.203}{2.014102 \times 931.3} \times 100 = 0.3840\,\%\,.$$

(*d*) In U^{235} fission, 200 MeV is released per uranium nucleus.

$$\therefore \left.\begin{array}{l}\text{The percentage of mass energy release in}\\ U^{235} \text{ fission}\end{array}\right\} = \frac{200}{235.0439 \times 931.3} \times 100 = 0.09137\,\%$$

$$\therefore \frac{\text{Energy release from } H^2 \text{ fusion}}{\text{Energy release from } U^{235} \text{ fission}} = \frac{0.3840}{0.09137} = 4.202.$$

35.10 Transuranic Elements

Elements with their atomic numbers greater than that of uranium (Z = 92) are called transuranic elements. All these are man-made and radioactive. Some of these elements are fissionable and hence useful. The following is the list of transuranic elements.

Z = 93	94	95	96	97	98
Np	*Pu*	*Am*	*Cm*	*Bk*	*Cf*
Neptunium	Plutonium	Americium	Curium	Berkelium	Californium
Z = 99	100	101	102	103	104
Es	*Fm*	*Md*	*No*	*Lw*	*Xv*
Einsteinium	Fermium	Mendelevium	Nobelium	Lawrencium	Kurchatovium

Such transuranic elements may be produced in the laboratory by the bombardment of certain heavy nuclides with neutrons. We give below typical methods of production, the reactions involved and the radioactive dacays of two of these nuclides.

(1) Neptunium (Z = 93) : When $_{92}U^{238}$ is bombarded with slow energy neutrons, neptunium is formed according to the reaction

$$_{92}U^{238} + {}_0n^1 \rightarrow {}_{92}U^{239}$$

$$_{92}U^{239} \rightarrow {}_{93}Np^{239} + {}_{-1}e^0\,.$$

(2) Plutonium (Z = 94) : Neptunium ($_{93}Np^{239}$) is itself radioactive. It emits a β-particle and produces plutonium according to the reaction

$$_{93}Np^{239} \rightarrow {}_{94}Pu^{239} + {}_{-1}e^0$$

Plutonium emits α-particles and decays into $_{92}U^{235}$ with a half-life of 24000 years.

$$_{94}Pu^{239} \rightarrow {}_{92}U^{235} + {}_2He^4\,.$$

EXERCISE

1. Write an essay on nuclear fission and its important applications.
2. Distinguish between nuclear fission and fusion. Explain the principle of a nuclear reactor. Mention some of its uses.
3. Describe the construction and working of a nuclear reactor. When is the reactor said to be critical ?
4. Give an account of the discovery and properties of transuranic elements.

5. A reactor is developing energy at the rate of 32×10^6 watts. How many atoms of *U*–235 undergo fission per second ? Assume that on the average, an energy of 200 MeV is released per fission.

[**Hint.** Energy released per second = 32×10^6 J. Energy released per fission = 200 MeV

$= 200 \times 10^6 \times 1.6 \times 10^{-19}$ J $= 3.2 \times 10^{-11}$ J.

$$\text{No. of atoms undergoing fission per second} = \frac{32 \times 10^6}{3.2 \times 10^{-11}} = 10^{18}\ .]$$

6. A city requires 100 Megawatts of electrical power on the average. If this is to be supplied by a nuclear reactor of efficiency 29% using *U*-235 as a nuclear fuel, calculate the amount of fuel required for 1 day's operation. Given that the energy released per fission of *U*-235 nuclide is 200 MeV.

Ans. 0.527 kg.]

7. Determine the Q value for the fission reaction

$_{29}Cu^{63} + {_1}H^1 \to {_{11}}Na^{24} + {_{19}}K^{39} + {_0}n^1$ [**Ans.** –24.15 MeV]

8. At what rate will $_{92}U^{235}$ metal be consumed by a reactor, operating at an output of 100 kW?

[**Ans.** 8.47×10^{-6} *kg/h*]

9. Given that the mass of the sun is 1.98×10^{50} kg and the solar energy is by fusion of 5.64×10^{11} kg of hydrogen to helium per second, calculate (*i*) the energy release in watt and (*ii*) the time to lapse for the sun to reduce its mass by 1%. [**Ans.** 3.7×10^{25} watt; 1.28×10^{11} years]

10. The fusion reaction $2\,{_1}H^2 \to {_2}He^4 + \text{energy}$, is proposed to be used for the production of industrial power. Assuming the efficiency of the process to be 30%, find how many kg of deuterium will be consumed in a day for an output of 50,000 kW. Given : Mass of $_1H^2 = 2.014102\ u$. Mass of $_2He^4 = 4.002604\ u$.

[**Hint.** $\Delta m = 0.025600\ u$. Energy released = 23.84 MeV. Since the efficiency of the process is 30%, the actual output is 7.152 MeV. ∴ Actual output per deuterium atom = 7.152/2 = 3.576 MeV = 3.576 × (1.6×10^{-13}) J. The output required = 5×10^7 J. ∴ No. of deuterium atoms required to produce the output required = 8.738×10^{19}. The mass of 8.738×10^{19} atoms = 2.923×10^{-7} kg. This is the consumption of deuterium in one second. ∴ The consumption of deuterium in one day is 2.525×10^{-2} kg.]

11. The fission fragments in the thermal neutron fission of U^{235} are found to be $_{42}Mo^{98}$ and $_{54}Xe^{136}$. What are the elementary particles released in the reaction ? Estimate the energy liberated in the reaction, given the masses of U^{235}, Xe^{136} and Mo^{98} are 235.044, 135.907 and 97.906 *u* respectively.

[**Ans.** 2 neutrons and 4 electrons; 205.06 MeV.]

12. What is the energy released when 1 kg of nuclear fuel is consumed if the fusion reaction $_1H^2 + {_1}H^2 \to {_2}He^4$ is possible ? [**Ans.** 5.75×10^{14} J]

13. Why are neutrons moderated to thermal speeds in nuclear reactors ?

[**Ans.** Slow neutrons have a greater probability of colliding against a nucleus than fast ones so as to cause fission]

14. What are the differences and similarities between nuclear fission and nuclear fusion ?

[**Ans.** In fission, a large nucleus splits into smaller ones. In fusion, two or more light nuclei combine together to form a single heavy nucleus. In both processes, there is a decrease in mass in the nuclear reaction. The decrease in mass is converted into energy, according to Einstein's equation $E = mc^2$.]

15. Why is it possible to produce the fission of U^{235} with slow neutrons whereas it is necessary to use fast neutrons to produce the fission of U^{238} ?

[**Ans.** When a neutron is captured by one of these nuclei, an energy equal to the corresponding activation energy must be supplied if fission is to occur. The actual energy supplied is equal to the sum of the kinetic energy and the binding energy of the neutron.

For U^{235}, the binding energy released by the captured neutron in forming U^{236} is 6.8 MeV. This is *greater* than the required 6.4 MeV of *activation* energy. Hence fission would occur even if the kinetic energy of the incident neutron is near zero. This is why thermal neutrons (energy 0.03 eV) cause fission in U^{235}.

For U^{238}, however, the B.E. released by the captured neutron in U^{239} is only 5.5 MeV which is *less* than the required 6.6 MeV of *activation* energy by 1.1 MeV. The incident neutron must therefore have a K.E., of atleast 1.1 MeV to induce fission. Hence, fast neutrons with energy above 1 MeV only, can cause fission in U^{238}.]

CHAPTER 36

NUCLEAR REACTORS

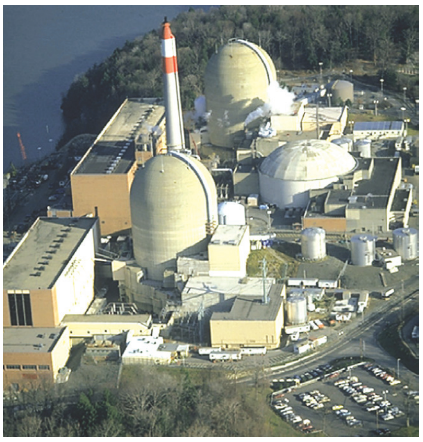
AT A GLANCE

36.1 Pressurized Water Reactor (PWR)
36.2 Boiling Water Reactor (BWR)
36.3 Fast Breeder Reactor
36.4 Neutron Cycle in a Thermal Nuclear Reactor
36.5 Fusion Reactors
36.6 Plasma Confinement

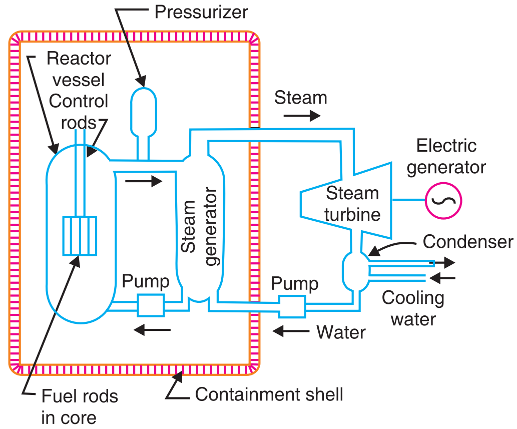
36.1 Pressurized Water Reactor (PWR)

If the pressure on the water surface is increased, its boiling point increases so that it can have more quantity of heat energy per unit mass. This principle is utilised in PWR.

Fig. 36.1 shows a simple sketch of a PWR. The fuel is in the form of uranium oxide (UO_2). It is sealed in long, thin zirconium-alloy tubes that are assembled together with movable control rods into a core. The core is enclosed in a steel pressure vessel. Light water is used as coolant and moderator.



Pressurized Water Nuclear Reactor.



Pressurizer
Reactor vessel
Control rods
Steam
Electric generator
Steam generator
Steam turbine
Condenser
Pump
Pump
Cooling water
Water
Fuel rods in core
Containment shell

Fig. 36.1

The water that circulates past the core is kept at a sufficiently high pressure, about 150 atm, to prevent boiling. The water enters the pressure vessel at about 280º C and leaves at about 320º C. The water then passes through a heat exchanger (steam generator) which produces steam. The resulting steam then passes out of the containment shell. The containment shell serves as a barrier to protect the outside world from accidents to the reactor. The high pressure steam drives a turbine. The turbine drives an electric generator and produces electrical power. Steam after running the turbine gets condensed into water in the condenser. This water is again circulated through heat exchanger by means of a pump.

36.2 Boiling Water Reactor (BWR)

Uranium is used as fuel in BWR (Fig. 36.2). Water is used as a coolant and moderator. Circulating water absorbs heat produced by fission of the fuel and gets converted into high pressure steam in the reactor core itself. This steam runs the turbine which in turn produces electric power. Steam after running the turbine gets condensed into water in the condenser. This water is again circulated into the reactor core by using a pump. The steam leaving the reactor may be slightly radioactive. Therefore, the pipe lines and turbine assembly are properly shielded.

Boiling Water Reactor.

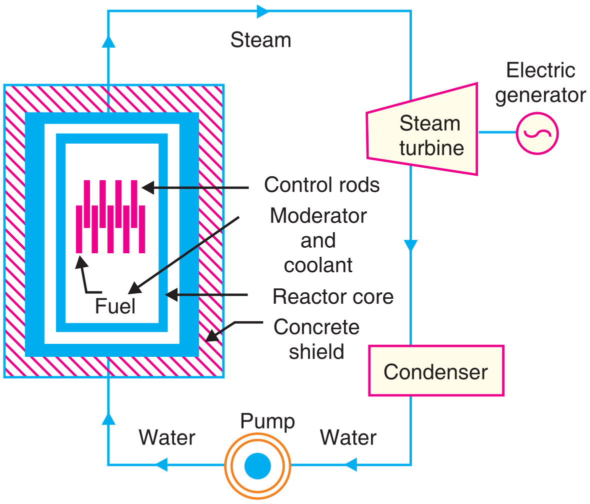
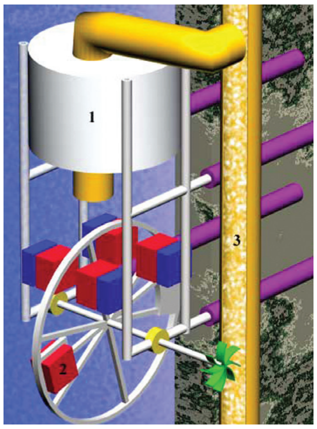


Fig. 36.2

36.3 Fast Breeder Reactor



Fast Breeder Reactor.

It is possible to run a reactor on fast neutrons by using a fuel with $^{239}_{94}$Pu. The fast neutrons released in the fission are used to transform $^{238}_{94}$U into $^{239}_{92}$Pu. A fast breeder reactor is one designed to function with fissions induced by fast neutrons. The core of such a reactor will contain two materials, say, fissionable fuel $^{239}_{94}$Pu and the potential fuel $^{238}_{92}$U. Existing breeder reactors are all of the fast type. They use a mixture of UO_2 and PuO_2 as fuel. The coolant is liquid sodium.

Fig. 36.3 shows a liquid-sodium-cooled fast breeder reactor. The core may consist of ^{235}U and ^{239}Pu, while the blanket contains the fertile ^{238}U that will breed into fissionable material. Liquid sodium is very effective as a medium for heat transfer. But its chemical reactivity and the radioactivity induced in it by neutron irradiation pose serious problems. To prevent contaminating the steam-generating system in the event of a sodium leak, an intermediate liquid sodium cooling loop is used between the primary loop that passes through the reactor core and the steam generator.

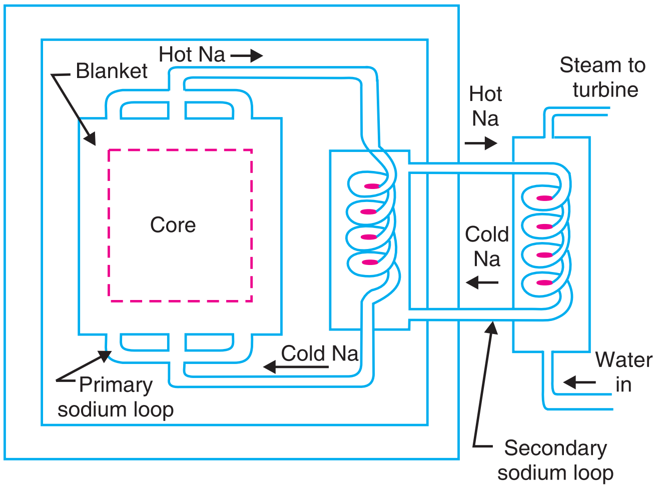


Fig. 36.3

36.4 Neutron Cycle in a Thermal Nuclear Reactor

Let us study the behaviour of neutrons in a reactor assembly consisting of enriched uranium (a mixture of the uranium isotopes $^{235}_{92}U$ and $^{238}_{92}U$) and a moderator. For simplicity we assume that the reactor is infinitely large so that there is no leakage of neutrons through its surface. Let us start with the fission of a $^{235}_{92}U$ nucleus by a thermal neutron. In this fission, the number of fast neutrons produced is ν. Of these, some will cause a fission in $^{238}_{92}U$, and therefore create additional fast neutrons. We take account of this small increase in the neutrons by introducing a factor, namely, ε , called the *fast fission factor.* The number of fast neutrons available is now $\nu\varepsilon$, where $\varepsilon > 1$. ε is usually about 1.03.

These $\nu\varepsilon$ neutrons diffuse through the pile and are slowed down by collisions with moderator nuclei. However, a few of them are captured by $^{238}_{92}U$ before they are slowed down to thermal energies. This is known as resonance absorption. Therefore, we define *resonance escape probability*, p, as the fraction of neutrons escaping resonance absorption. Hence, the number of neutrons surviving thus far and reaching thermal energies is $\nu\varepsilon p$. p is usually 0.95.

Of these $\nu\varepsilon p$ thermal neutrons, a fraction f called the *thermal utilisation factor* is absorbed in $^{235}_{92}U$ nuclei while the rest are absorbed in moderator and other structural materials. f is always less than 1. Thus the number of $^{235}_{92}U$ nuclei undergoing fission is $\nu\varepsilon p f$.

When these neutrons are absorbed by $^{235}_{92}U$ in the fuel, some cause fission, whereas some produce other kinds of reactions such as capture. The fraction of neutrons causing fission is given by the ratio of fission cross-section (σ_f) to the total absorption cross-section (σ_a).

$$\therefore \text{ the number of second generation thermal neutrons causing fission of } ^{235}_{92}U = \nu\varepsilon p f \frac{\sigma_f}{\sigma_a}.$$

Now, $\nu \dfrac{\sigma_f}{\sigma_a} = \eta$ = *the average number of fission neutrons released per neutron absorbed by a fissionable nuclide.*

The ratio K of the number of fissions produced to a single neutron is

$$K = \eta \varepsilon p f.$$

This is the *multiplication factor*. We have assumed that the reactor assembly is of an infinite size so that there is no leakage of neutrons. Hence

$$K_\infty = \eta \varepsilon p f.$$

This relation is called the *four-factor formula*. In designing a nuclear reactor, the fundamental problem is to maintain K at 1 in order to achieve a self-sustaining chain reaction. Of the four factors, η and ε are fixed for a given fuel. The other two factors p and f are made as large as possible by properly designing the reactor geometry, arranging the fuel, and choosing the moderator.

36.5 Fusion Reactors

In the experiments conducted by bombarding the particles ${}_1^2\text{H}$ and ${}_1^3\text{H}$ with the deuterons, the following reactions have been found to occur:

$${}_1^2\text{H} + {}_1^2\text{H} \rightarrow {}_1^3\text{H} + {}_1^1\text{H} + 4.0\,\text{MeV} \qquad \text{...(1)}$$

$${}_1^2\text{H} + {}_1^2\text{H} \rightarrow {}_2^3\text{He} + {}_0^1n + 3.3\,\text{MeV} \qquad \text{...(2)}$$

$${}_1^3\text{H} + {}_1^2\text{H} \rightarrow {}_2^4\text{He} + {}_0^1n + 17.6\,\text{MeV} \qquad \text{...(3)}$$

Each of the deuteron-deuteron reactions (1) and (2) has a small cross-section. However, the deuteron-triton reaction, (3) is much more probable and also liberates a large amount of energy. The first fusion reactors are more likely to employ a deuterium—tritium mixture.

Three conditions must be met by a successful fusion reactor.

1. We know that for a fusion reaction to take place, the colliding particles should have enough energies to overcome the coulombian repulsive forces between them. This needs energies of the order of 0.1 MeV which corresponds to a temperature of approximately 10^8 K. This means that for a fusion reaction to take place temperatures of this order should prevail initially. At such high temperatures, the atoms are ionized into positively charged ions (*i.e.,* nuclei) and electrons. This completely ionised and electrically neutral state of matter is called *plasma.*

2. There should be a fairly high density of the nuclei to ensure that collisions between nuclei are frequent. An important parameter for the fusion reaction is the *reaction rate* which can be represented as

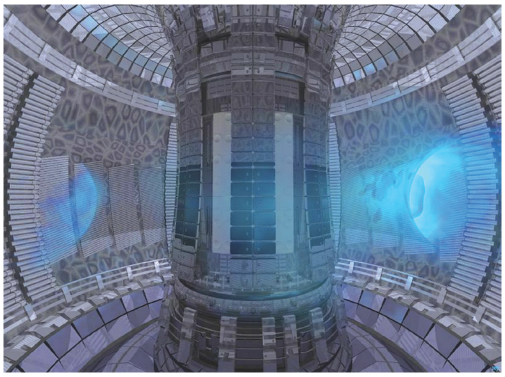
$$R_{12} = n_1 n_2 < \sigma v_{12} > \text{reactions/m}^2\text{/sec.}$$

Here n_1, n_2 represent particle densities of the colliding particles, v_{12}, the relative velocity of the particles and $<\sigma v_{12}>$ represents the average of the product of cross-section σ and the relative velocity of the particles.

If the colliding particles are identical, we have

$$R_{11} = \frac{1}{2} n^2 < \sigma v > \text{reactions/m}^2\text{/sec.}$$

3. The assembly of reacting nuclei must remain together for a long enough time to give off more energy than the reactor's operation takes.



Fusion Reactor.

It is difficult to achieve a high enough particle density n for a long enough time τ in a hot plasma to obtain a net energy yield. According to the *Lawson criterion*, the product of the density and the confinement time (*i.e.*, the product $n\tau$) of the plasma should be about 2×10^{22} particles sm^{-3} and 10^{20} particles sm^{-3} for the deuteron-deuteron and the deuteron-triton fusion reactions respectively to occur in it.

36.6 Plasma Confinement

(*i*) Magnetic confinement. For the confinement of the plasma, one cannot use the walls of any vessel. Any contact with the wall will not only quickly cool the plasma but also cause the wall to evaporate. There are at present two schemes under investigation for confining the thermonuclear fuel: (*i*) *magnetic confinement* and (*ii*) *inertial confinement*.

In magnetic confinement, the plasma is confined by a carefully designed magnetic field. Two classes of field geometry are being studied. They are: (1) Magnetic bottle and (2) Tokamak.

1. Magnetic bottle. It is based on the reflection of an ion that moves in a magnetic field B whose lines of force converge. The magnetic force on an ion of velocity v is perpendicular to both v and B. So the force has a backward component in such a field as well as the inward component that causes the ion to follow a helical path around the lines of force (Fig. 36.4). If the backward force is powerful enough, the converging field acts as a mirror to reverse the ion's direction of motion. Thus, we can form a high density of magnetic field lines which reflects the particles back into the low-field region and is hence known as *a magnetic mirror*. A pair of magnetic mirrors constitutes a magnetic bottle (Fig. 36.5).

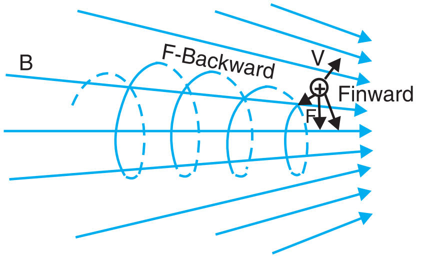


Fig. 36.4

2. Tokamak. A 'tokamak' is a magnetic bottle in which the confining field is created by a combination of currents flowing in external coils and currents flowing in the plasma. It is a toroidal magnetic trap.

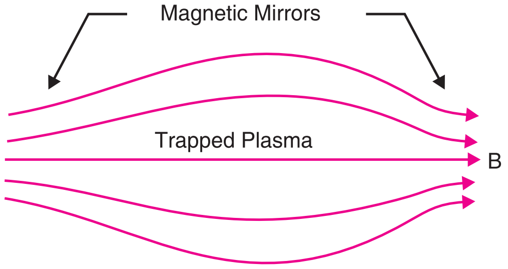


Fig. 36.5

The principle of tokamak method of magnetic confinement is shown in Fig. 36.6. A toroidal field is produced by a winding of coils, and a poloidal field is produced by an axial current. These two fields are combined in the tokamak design. (The current-carrying coils are not shown). The resulting field lines form a helix, through which the ions can travel in closed orbits.

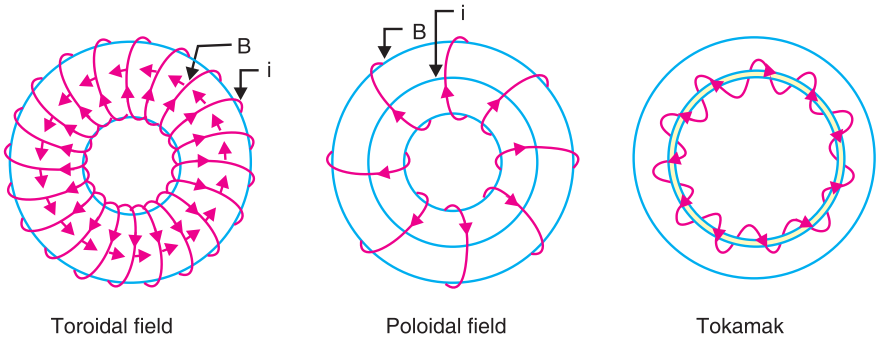


Fig. 36.6

***(ii)* Inertial confinement.** This scheme employs energetic beams to both heat and compress tiny deuterium-tritium pellets. Here, the fusion fuel (*e.g.*, mixture of deuterium and tritium), in the form of a pellet, is imploded from all sides by energy sources such as laser beams, high energy electron or ion beams (Fig. 36.7). The intense compression pressures and the high temperatures produced in the pellet may produce conditions conducive to fusion. The difficulties in this approach are the low efficiencies of laser or other sources, and the need to produce stable symmetrical implosion.

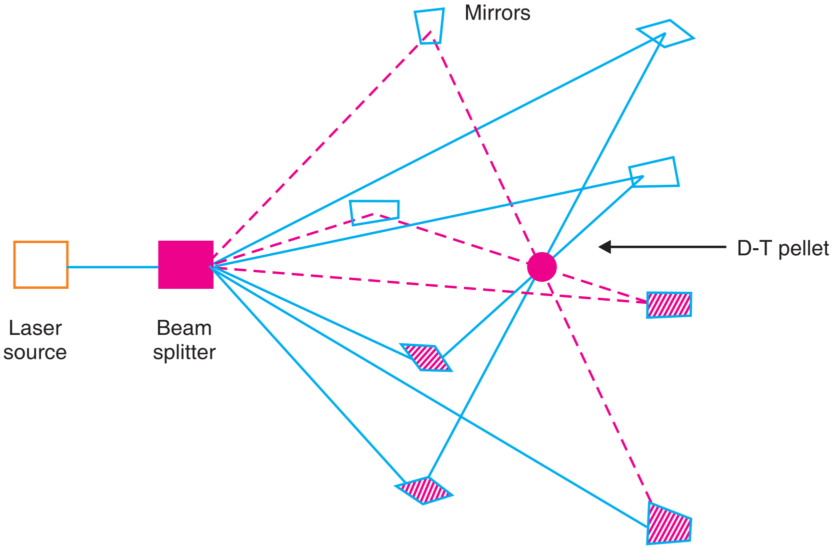
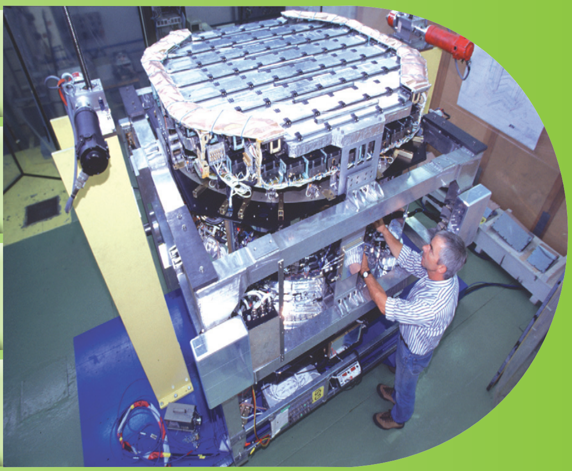


Fig. 36.7

Both the above schemes are in advanced experimental stage. The experiments have shown that fusion reactions can be initiated but it has not been possible to extract the useful amounts of energy.

Fusion reactors have two important advantages over fission reactors as sources of mechanical and electrical power.

1. The fuel for fusion reactors (Deuterium and tritium) is much more abundant on earth than that for fission reactors (Uranium or Thorium).
2. Fusion reactors do not produce large amounts of radioactive residues which are the usual by products of the operation of fission reactors.



CHAPTER 37

COSMIC RAYS AND THE UNIVERSE

AT A GLANCE

37.1	Discovery of Cosmic Rays	**37.2**	Latitude Effect
37.3	The East-West Effect (Azimuth Effect)	**37.4**	Altitude Effect
37.5	Primary Cosmic Rays	**37.6**	Secondary Cosmic Rays
37.7	Cosmic Ray Showers	**37.8**	Discovery of Positron
37.9	The Mesons	**37.10**	Van Allen belts
37.11	Origin of Cosmic Rays	**37.12**	The Big Bang Theory
37.13	Thermal History of the Universe	**37.14**	Hubble's Law
37.15	The Future of the Universe		

37.1 Discovery of Cosmic Rays

Cosmic rays are highly penetrating radiations which are continually entering the earth's atmosphere in all directions from outer space. Cosmic rays consist of high energy charged particles. Most of these particles have energy of the order of 15 GeV.

Portable Cosmic Ray Detector.

Cosmic rays were first detected by Elster and Geital (1899) and by C.T.R. Wilson (1900). They found that the charge of a well insulated electroscope leaked away slowly even in the absence of any ionizing agent. At first, it was thought that the loss of charge of electroscope is due to the ionizing radiations coming from *radioactive substances* present in the earth. In that case, the rate of discharge must decrease when the electroscope is taken to a height above the earth. Hess in 1911, recorded ionization of air at different altitudes by sending eletroscopes in balloons and found that the rate of discharge from electroscope increases as we go higher above the earth. On the other hand, the ionization of air in the electroscope decreases as we go down into mines or the sea.

From these observations, it was concluded that *some kind of penetrating rays were entering the earth's atmosphere from outer space in all directions.* These were named as cosmic rays by Millikan.

37.2 Latitude Effect

The variation of cosmic ray intensity with latitude is shown in Fig. 37.1. The cosmic ray intensity is maximum at the geomagnetic poles ($\lambda = 90°$) and minimum at the geomagnetic equator ($\lambda = 0$). The intensity remains constant between 42° and 90°. This variation of cosmic ray intensity with geomagnetic latitude is called latitude effect.

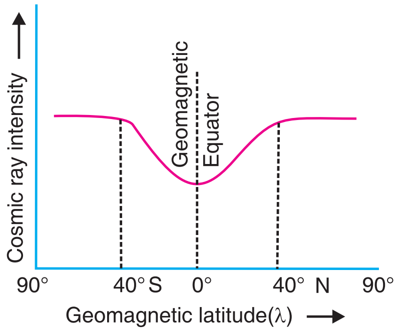


Fig. 37.1

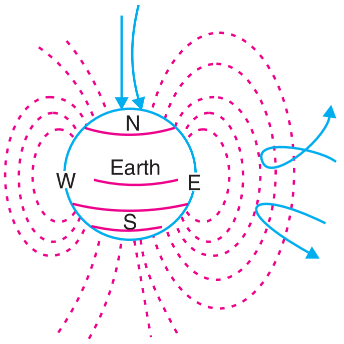


Fig. 37.2

Explanation of latitude effect. The latitude effect may be explained on the basis of *effect of earth's magnetic field on cosmic rays. The presence of such an effect indicates that the cosmic rays are charged particles*. The earth's magnetic field is directed from south to north. The earth's magnetic field at the equator is perpendicular to the direction of travel of charged cosmic ray particles. Therefore, it exerts maximum force ($F = qv\,B \sin 90°$) upon the particles which consequently suffer maximum deflection. They are deflected away from the earth (Fig. 37.2). The intensity of cosmic rays is therefore, minimum at the equator. At poles, cosmic ray particles move parallel to the earth's field. They, therefore, suffer minimum deflection. Hence the intensity of cosmic rays is maximum at the poles.

Mathematical analysis shows that for a given latitude (λ), there is a minimum momentum (p_{min}) below which no particle can reach the earth. The minimum momentum p_{min} at zenith is given by

$$p_{\text{min}} = 15 \cos^4 \lambda \text{ (Be V/c)}$$

37.3 The East-West Effect (Azimuth Effect)

It is found that the *number of cosmic ray particles coming from the west is greater than those coming from the east.* This effect is called east-west effect and is maximum at the equator. At the equator, the number of particles coming from west is 14% more than the particles coming from east. This phenomenon gave an evidence of the fact that *cosmic rays are composed predominantly of positively charged particles.*

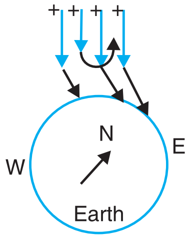


Fig. 37.3

Explanation : The charged particles approaching the earth's atmosphere are deflected by the earth's magnetic field in a direction perpendicular to the magnetic field and to the direction of their motion. According to Fleming's left hand rule, the positively charged particles are deflected towards the east by earth's magnetic field and thus appear to come from the west of the vertical (Fig. 37.3). Thus at any azimuth, more particles approach the earth from the west than from the east.

If I_W and I_E are the intensities of cosmic rays coming from the west and east respectively, then the east-west asymmetry is defined as

$$\frac{I_W - I_E}{[(I_W + I_E)/2]}.$$

37.4 Altitude Effect

The intensity of cosmic rays was measured by measuring the ionization produced by cosmic rays at different altitudes. The variation of cosmic ray intensity with altitude is shown in Fig. 37.4. It is observed that the intensity rises slowly upto a height of about 8 km, after which the rise becomes fast upto about 19 to 24 km. At heights above 24 km, the intensity starts decreasing gradually. The experiments were conducted at 3°, 38° and 51° N and the results are similar.

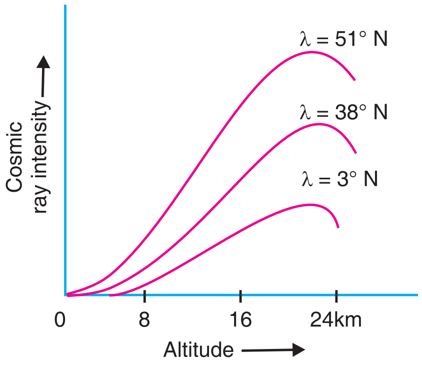


Fig. 37.4

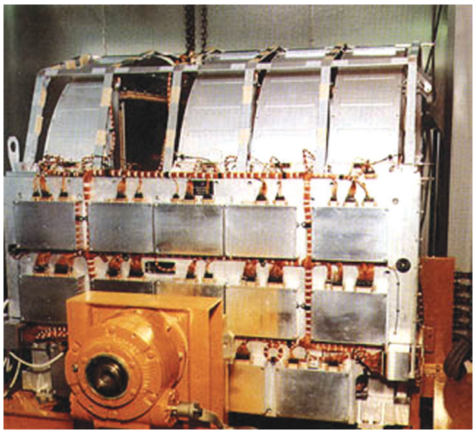
The maxima of intensity is not at the top of the atmosphere but well below it. The reason is that at this height the primary cosmic rays produce quite a good number of secondary particles due to interaction with nuclei of atmospheric gases. Thus both primary and secondary rays are present in abundance at that height. With decrease of altitude the absorption increases and, therefore, intensity falls.

Longitude Effect

The intensity of cosmic rays also depends upon the longitude of the point of observation. It is called longitude effect. The intensity of cosmic radiation along the equator varies at different longitudes. This equatorial variation is attributed to the fact that the earth's magnetic field is not symmetrical about its axis.

37.5 Primary Cosmic Rays

The cosmic rays which are just entering our earth's atmosphere from outer space are called primary cosmic rays. Primary cosmic rays consist mainly of positively charged atomic nuclei with Z upto about 40. About 90% of the primaries are protons, 9% helium nuclei and the remaining heavy nuclei. The energies of primary cosmic rays range from 1 MeV to 10^{14} MeV.



Primary Cosmic Rays Experimental Setup.

37.6 Secondary Cosmic Rays

When primary cosmic rays interact with the nuclei of atmospheric gases, secondary cosmic rays are produced. Below an altitude of 20 km, all cosmic radiation is secondary. On entering the atmosphere, the primary cosmic rays collide with air nuclei and produce mostly π-mesons (positive, negative and neutral) and some hyperons. The π-mesons so produced carry sufficient energy and decay into lighter particles : μ-mesons, electrons, positrons, neutrinos and photons. All such particles constitute the secondary cosmic rays. At sea level the secondary cosmic rays contain nearly 70% μ–mesons, 29% electron-positron pairs and 1% heavy particles. The mesons in the secondary cosmic rays constitute the *hard* component and the electrons, positrons and photons constitute the *soft* component.

37.7 Cosmic Ray Showers

Cosmic ray shower is the name given to a grouping of cosmic ray particles that have been produced by some common cause and not by mere chance.

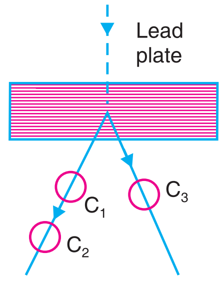
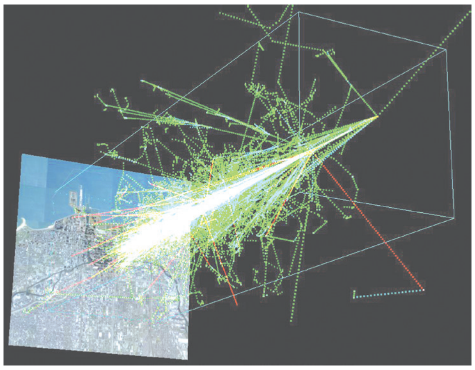


Fig. 37.5

Rossi investigated the phenomenon by using an arrangement of three coincidence counters in a triangular pattern with a lead plate above the counters (Fig. 37.5). A coincidence discharge of the three counters can be produced only by the simultaneous passage of at least two particles, including the incident one. Rossi observed that an appreciable number of coincidence counts were registered. He found that the number of coincidences increased as the thickness of the lead plate was increased. From this result he inferred that the three counters were affected by two or more secondary particles produced simultaneously by a single cosmic ray as it penetrated the lead plate. Thus the production of showers in the lead plate was detected. Rossi's results were explained by Bhabha and Heitler and, independently, by Carlson and Oppenheimer. They gave the cascade theory of cosmic ray showers.



Cosmic Ray Showers.

Cascade theory of showers. Shower production involves two processes, *viz.*, *radiative collision* and *pair production*.

A high energy electron (or positron), present in cosmic rays, loses energy, when it encounters the atomic nuclei in earth's atmosphere. The energy appears as high energy photon. The photon interacts with the electric field of an atomic nucleus and is completely absorbed, resulting in the production of an electron-positron pair. The energy required for pair production is more than 1 MeV. The electron and positron so produced have sufficient energy to produce more photons on interaction with nuclei. These photons are further capable of bringing about pair production. The result is the generation of a large number of photons, electrons and positrons having a common origin (Fig. 37.6). The multiplication will continue until the initial energy becomes divided between a large number of pairs and the individual energies of the particles fall below the "critical energy" when photon emission and pair-production can no longer occur.

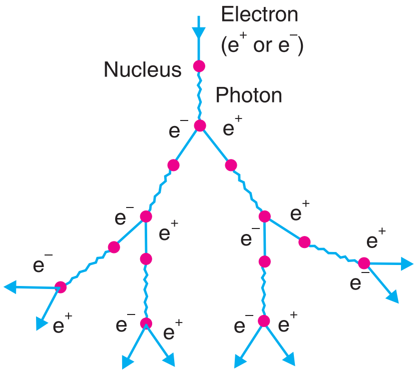


Fig. 37.6

Pair Production. The conversion of a *photon* into an *electron* and a *positron*, when the photon traverses the strong electric field surrounding a nucleus, is called pair production. The electric charge is conserved, since the electron and positron have charges of equal magnitude and opposite sign. The energy of the photon provides :

(*i*) the rest energy of the electron ($m_0 c^2$)
(*ii*) the rest energy of the positron ($m_0 c^2$)
(*iii*) the K.E. of the electron (E^-) and
(*iv*) the K.E. of the positron (E^+)

If $h\nu$ is the energy content of the radiation, then

$$h\nu = 2m_0 c^2 + E^+ + E^-$$

Pair production is *materialization of radiant energy*. It is clear from the equation that a photon cannot create an electron-positron pair unless its energy exceeds $2m_0 c^2$. This gives the *threshold energy* for pair production.

Annihilation of Matter. The converse of the materialization of energy (production of electron-positron pair by photons) is the annihilation of matter. When a positron combines with an electron both disappear producing two quanta of γ rays: [$e^+ + e^- \rightarrow 2\gamma$]. This process is called annihilation of matter.

37.8 Discovery of Positron

The positron is the *exact counterpart of the electron*, having the same mass and having one unit of positive charge. It was first discovered by Anderson in 1932.

Anderson was making cosmic ray investigations with a cloud chamber placed in a strong magnetic field. In one of the photographs, he observed two tracks of the same curvature but in opposite directions (Fig. 37.7). The curve *A* was identified as due to electron. The track *B* was at first thought to be due to an electron moving from above downwards. But measurements of the energy of the particles on the two sides of the lead plate showed that the energy above the plate was less than that below the plate. This showed that the particle is moving from below upwards, just like the electron causing the track *A*. It was, therefore, deduced that the track *B* was caused by a particle of the same mass as the electron, but with a *positive* charge *e*. The particle was called a *positron*.

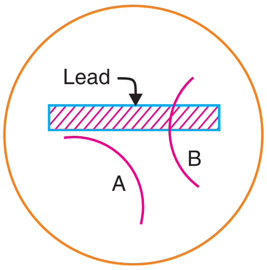


Fig. 37.7

37.9 The Mesons

The meson is a fundamental particle, possessing a mass intermediate between the electron and proton. The known varieties of mesons are:

(*i*) π-mesons or pions (*ii*) *K*-mesons or kaons (*iii*) η-mesons or eta.

(*i*) π-mesons or Pions: π-mesons were discovered in 1947 in the cosmic rays. They can exist in three states: π^+, π^- and π^0. The π^+ and π^- are antiparticles of each other. π^+ and π^- mesons have a rest mass of 273 m_e. The rest mass of π^0 meson is 264 m_e.

The pion is unstable with a mean life of 2.6×10^{-8} *s*. The charged pions decay into corresponding muon and a μ-neutrino :

$$\pi^+ \rightarrow \mu^+ + v_\mu$$

$$\pi^- \rightarrow \mu^- + \overline{v_\mu}$$

The neutral pion has a mean life of 8.7×10^{-17} s and decays into two gamma rays :

$$\pi^\circ \rightarrow \gamma + \gamma$$

μ^+ and μ^- (positive and negative muons) have the same rest mass of 106 MeV $/c^2$ (207 m_e) and the same spin of 1/2. Both decay with a relatively long mean life of 2.2×10^{-6} s into electrons and neutrino-antineutrino pairs.

$$\mu^+ \to e^+ + \nu_e + \overline{\nu_\mu}$$

$$\mu^- \to e^- + \nu_\mu + \overline{\nu_e}$$

(*ii*) *K*-Mesons (or Kaons): *K-meson is a heavier class of mesons.* $\boldsymbol{K}^+$ has a rest mass of 966 m_e and mean life of 1.2×10^{-8} s. There are two varieties of neutral *K*-mesons → K_1^0 and K_2^0. Both of them have rest masses of 974 m_e but their mean lives are 9×10^{-11} s and 5×10^{-8} s respectively.

(*iii*) η-Meson: The neutral eta-meson (η^0) was discovered in 1961. It has a rest mass of 1073 m_e and a zero spin value (boson). Its mean life is 6×10^{-19} s.

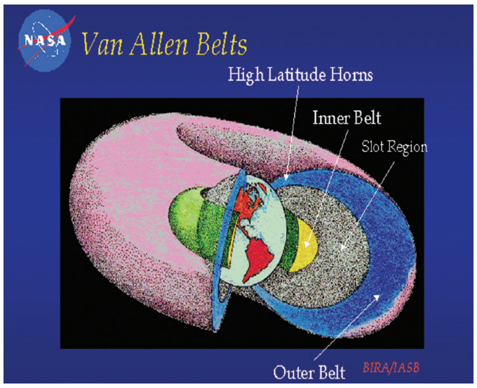
37.10 Van Allen Belts

Artificial satellites helped to discover the presence of two strong radiation belts, called Van Allen belts. These belts are formed around the earth except at magnetic poles, at heights of 3200 km and 16000 km above the equator. Each belt contains two localized zones:

1. Inner Zone: It consists of protons of high energy of the order of 100 MeV and electrons of low energy of the order of 1 MeV.

2. Outer Zone: It is less intense than the inner zone and consists of only low energy electrons, the energy being of the order 0.1 MeV.

The Van Allen belts consist of high energy protons and electrons originating from the sun and cosmic rays that have been captured by the earth's magnetic field. The existence of such a belt proved beyond doubt that high intensity particles are trapped by the earth's magnetic field. Radiation belts are present around all celestial bodies which have a magnetic field. The moon has no magnetic field of its own and no radiation belt.



Van Allen Belts.

37.11 Origin of Cosmic Rays

Several theories have been put forward regarding the origin of *primary* cosmic rays, and the mechanism by which they acquire enormous energies.

(*i*) Explosion theory. Lemaitre, Regener and others have suggested that once upon a time the whole mass of the universe was concentrated in a single nucleus. About three billion years ago, this universe *exploded*, forming the galaxies which are still running away from each other. During the explosion, a fantastically great amount of radiation was formed. Protons and other nuclei were shot out in all directions with all energies and in sufficient numbers. The cosmic rays are simply the debris, the dust of the explosion. This view does not explain the presence of the *heavy nuclei* in primary rays. Such a tremendous explosion should have broken up the *heavy nuclei* into their constituent particles, *viz.*, protons and neutrons.

(*ii*) Origin from sun. One of the views is that the *sun may be the source of at least some of the cosmic rays.* According to it, at the times of solar activity violent eruptions occur and ionized gases shoot out from the sun. Thus some of the protons in the sun acquire high energies and thrown out into interplanetary space. The points in favour of this assumption are that the cosmic ray intensity increases during solar flares and varies slightly with the rotation of the sun. But since the cosmic ray intensity remains almost uniform at all hours of the day and night, sun can not be thought to be responsible for the majority of the primary cosmic ray particles. It may be the source of the small fraction of the low energy primaries.

(*iii*) Origin from cosmic ray stars. Another view is that the *cosmic rays come from the so called cosmic ray stars* which are more active than the sun. Our galaxy has about 10^{11} stars including double stars, variable stars, novae and supernovae. All these may be the possible origins of cosmic rays.

The current view of the origin of cosmic rays is that the *sun emits low energy cosmic rays while high energy cosmic rays are emitted by cosmic ray stars within our galaxy.*

Regarding the *tremendous energies* of cosmic rays, most scientists believe that *protons acquire acceleration in interstellar magnetic fields.*

THE UNIVERSE

37.12 The Big Bang Theory

According to this theory, all matter in the universe was concentrated in a very dense and very hot primeval nucleus at the beginning. A cosmic explosion occurred some 15 billion years ago and the matter was thrown out in the form of galaxies. The galaxies thus formed are continuously moving away from each other. The distances between the galaxies will continually increase and space will grow more and more empty. The hydrogen in each galaxy will be gradually used up and the galaxies themselves will become fainter and fainter. Eventually the galaxies will fade out completely. In the end, the universe will be devoid of all matter, energy and life.

We have three observations that strongly support Big-Bang cosmology:

1. The uniform expansion of the universe.
2. The relative abundance of hydrogen and helium in the universe.
3. The cosmic microwave background radiation.

37.13 Thermal History of the Universe

Cosmologist are attempting to reconstruct the story of *Big Bang* but cannot yet do that back to time zero. Nothing can be said about the state of the universe until 10^{-43} s after the Big Bang in the absence of a quantum-mechanical theory of gravity. We shall discuss what the universe must have been like from about 10^{-43} s after the Big Bang to the present time.

As the initial compact, intensely hot fireball of matter and radiation from the Big Bang expanded, it cooled and underwent a series of transitions at specific temperatures. Fig. 37.8 illustrates the time-temperature relationship in the universe. The unit of kT here is eV. To convert temperature T in kelvin to kT in electron-volts, the formula is

$$kT\,(\text{eV}) = 8.6 \times 10^{-5}\, T\,(\text{K}).$$

The initial expansion began about 2×10^{10} years ago. Cosmologists refer to this event as the *initial singularity* because at this time the volume of the Universe was zero and the density of mass-energy was infinite.

(*i*) From 10^{-43} to 10^{-35} s, the universe cooled from 10^{28} to 10^{23} eV. At energies like these the strong, electromagnetic and weak interactions are merged into a single interaction mediated by superheavy X bosons. At 10^{-35} s, particle energtiesbecame too low for free X bosons to be created any longer. The strong interaction became separated from the electroweak interaction. Quarks and leptons now became independent.

(*ii*) From 10^{-35} to 10^{-10} s, the universe consisted of a dense soup of quarks and leptons. The behaviour of quarks and leptons was controlled by the strong, electroweak, and gravitational interactions. At 10^{-10} s, the electroweak interaction became separated into the electromagnetic and weak components.

(*iii*) Around 10^{-6} s, the quarks condensed into hadrons.

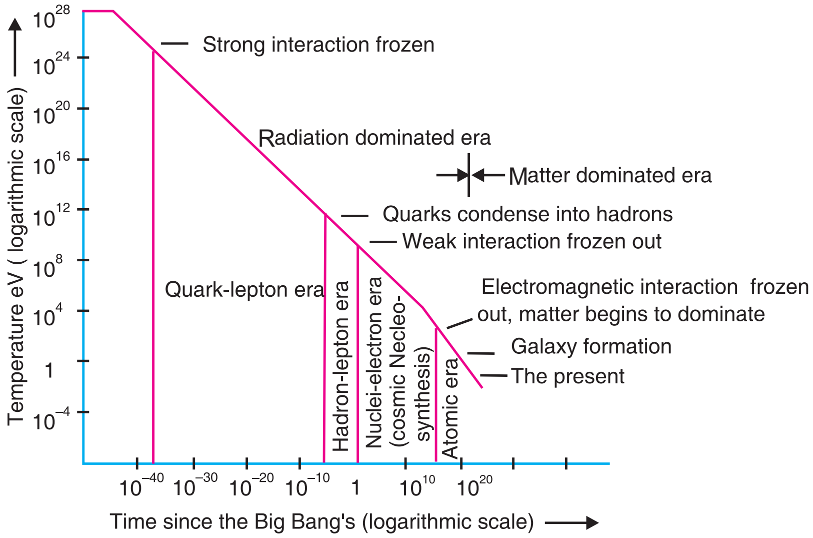


Fig. 37.8

(*iv*) At t = 1s, neutrino energies fell sufficiently for them to be unable to interact with the hadron-lepton soup. The neutrinos and antineutrinos that existed remained in the universe. But they did not participate any further in its evolution. From then on protons could no longer be transformed into neutrons by inverse beta-decay. But the free neutrons could beta-decay into protons:

$$n \to p + e^- + \bar{v}$$

As the temperature decreased, the number of neutrons fell while the number of protons rose. Some initial *nucleosynthesis* now takes place. Helium-producing nuclear reactions were starting to occur. Nuclear synthesis stopped at about t = 5 min. The small amount of nuclear matter is approximately 76% by mass protons and 24% α-particles. There are traces of deuterons and of $^3_2$He nuclei.

(*v*) From 5 min to around 100, 000 y after the Big Bang, the universe consisted of a plasma of hydrogen and helium nuclei and electrons in thermal equilibrium with radiation. Once the temperature fell below 13.6 eV, hydrogen atoms could form and not be disrupted. Now matter and radiation were decoupled and the universe became transparent. The electromagnetic interaction was frozen out.

After 2000 years, when matter (hydrogen and helium) began to dominate the Universe and radiation became a secondary constituent, the *matter era* began.

Note that the neutrinos decoupled at near 1s but the decoupling of the photons of the blackbody radiation from the surviving electrons, protons and the α-particles did not occur completely until the temperatures fell to about 10^4 K (0.86 eV) at an age of about 10^6 years. Until then Compton scattering and bremsstrahlung kept the photons in equilibrium with the electrons and nuclei. But, when the temperature fell well below the first excitation potentials of hydrogen and helium, the Universe became transparent to photons. Thereafter the blackbody radiation was free to cool independently of the matter.

(*vi*) The radiation left behind then continued to spread out with the rest of the universe, undergoing Doppler shifts to longer and longer wavelengths. The radiation cooled to become the 2.7 K cosmic blackbody radiation observed today.

(*vii*) Once matter and radiation were decoupled, gravity became the dominant influence on the evolution of the universe. The gaseous matter underwent a critically important transformation – it could now form the first galaxies. Prior to decoupling, radiation pressure kept the matter distribution smooth. Afterwards, small inhomogeneities could grow and condense into the first gravitationally bound systems.

37.14 Hubble's Law

It has been observed that the spectral lines of many elements in galaxies are shifted either towards the high frequency, called blue shift or towards the low frequency, called red shift, in relation to those observed in the laboratory. The Doppler shift indicates that the galaxies are either moving towards or away from the earth. Hubble found that galaxies in all directions are moving away from us and that there is a direct proportionality between the distance of a galaxy and its redshift. Hubble's law states that *the velocity of recession v of a galaxy is directly proportional to its distance from the earth* (R). Hence

$$v = HR.$$

H is known as Hubble's parameter. Thus we conclude that *the universe is expanding.*

37.15 The Future of the Universe

Will the universe continue to expand forever and will the material in it become cold, compact stellar objects? This depends on how much matter the universe contains and on how fast it is expanding. The expansion of the Universe can stop only if it contains sufficient mass for gravity to provide an adequate attractive force to slow down and reverse the expansion. The *critical density* ρ_c above which expansion will be arrested is about 6×10^{-27} kg m^{-3}, equivalent to about 3.5 hydrogen atoms per m^3. Let ρ be the average density of the universe. There are three possibilities.

(*i*) If $\rho < \rho_c$, the universe is *open* and the expansion will never stop (Fig. 37.9). Eventually new galaxies and stars will cease to form and existing ones will end up as black dwarfs, neutron stars, and black holes.

(*ii*) If $\rho > \rho_c$, the universe is *closed* and sooner or later gravity will stop the expansion. The universe will then begin to contract. The progression of events will be the reverse of those that took place after the Big Bang, with an ultimate Big Crunch – a fiery death. In that case perhaps the system will rebound in another big bang. A continuing cycle of big bang, expansion, contraction, big crunch is a possibility.

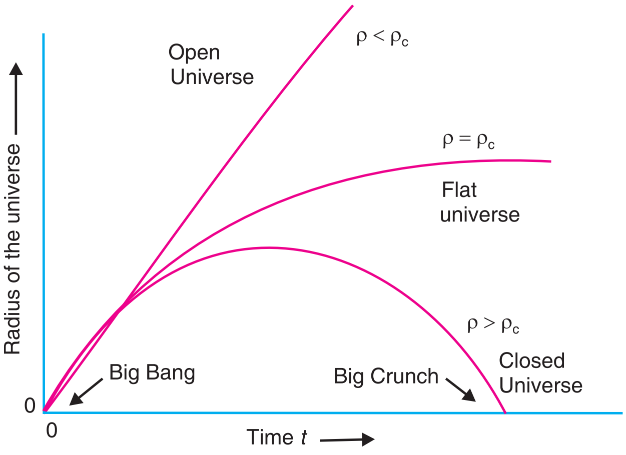


Fig. 37.9

(*iii*) If $\rho = \rho_c$, the expansion will continue at an ever-decreasing rate but the universe will not contract. In this case the universe is said to be *flat* because of the geometry of space in such a universe (Fig. 37.10). If $\rho < \rho_c$, space is negatively curved; a two-dimensional analogy is a saddle. If $\rho > \rho_c$, space is positively curved. The surface of a spherical ball has positive curvature. In all cases, however, spacetime is curved.

The flatness problem relates to the issue of whether or not the Universe's geometry is open or closed. We focus on this problem with the parameter Ω, where

Ω = (actual mass – energy density) / critical density.

If $\Omega = 1$, the geometry is flat; If $\Omega < 1$, it is hyperbolic; and if $\Omega > 1$, it is spherical and closed.

To find the value of the Critical Density ρ_c

Consider a spherical volume of the universe of radius R whose centre is the earth. Assume that the distribution of matter in the universe is uniform. Let ρ be the density of matter inside this volume.

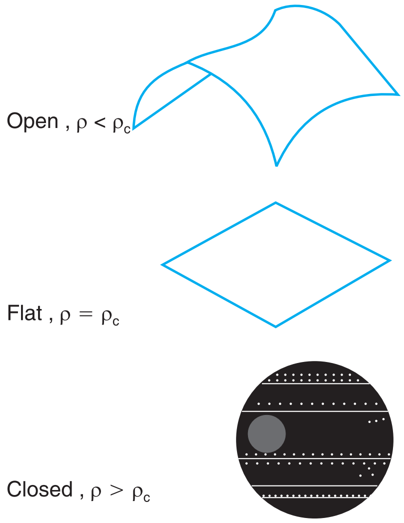


Fig. 37.10

Mass of the Universe $M = \frac{4}{3}\pi R^3 \rho.$

Let m be the mass of the galaxy.

Therefore, the total energy E of the galaxy is the sum of its kinetic and potential energies, *i.e.,*

$$E = \frac{1}{2}mv^2 - mMG/R .$$

To escape permanently from the universe the galaxy must have a minimum kinetic energy $\frac{1}{2}mv^2$ such that its total energy E is 0:

$$E = KE + U = \frac{1}{2}mv^2 - \frac{GmM}{R} = 0$$

$$\frac{1}{2}mv^2 = \frac{GmM}{R}$$

According to Hubble's law, $v = HR$, where H is *Hubbles's parameter.*

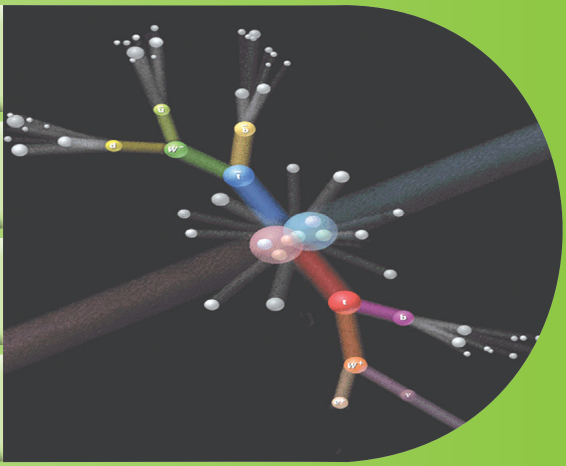
$$\therefore \qquad \frac{1}{2}m\,(HR)^2 = \frac{Gm}{R}\left(\frac{4}{3}\pi R^3 \rho_c\right)$$

Critical density $$\rho_c = \frac{3H^2}{8\pi G}.$$

If the density is greater than the critical value ρ_c, the galaxy is bound in the universe and will eventually fall back. On the other hand, if the density is less than the critical value, the galaxy will move away, and thus keep the universe expanding. Based on the astronomical observations, the present mass density has been estimated to be ~ 5×10^{-28} kg/m^3, about ten times less than the critical value. This seems to indicate that the universe will go on expanding for ever.

CHAPTER

38



ELEMENTARY PARTICLES I

AT A GLANCE

38.1	Introduction	38.2	Particles and Anti-particles
38.3	Antimatter	38.4	The Fundamental Interactions
38.5	Elementary-particle Quantum Numbers	38.6	Conservation Laws and Symmetry
38.7	The Quark Model		

38.1 Introduction

After studying the structure of atoms, one gets the impression that perhaps electron, proton and neutron are the only building blocks of matter. Recently the extensive studies made partly on high energy cosmic ray particles and even more, with the help of high energy accelerators, have revealed the existence of numerous new nuclear particles.

The subatomic or elementary particles discovered so far, form a long list (around 200). These particles are elementary in the sense that they are structureless, *i.e.*, they cannot be explained as a system of other elementary particles. Table 38.1 shows the elementary particles having relatively long life. Their classification is as follows :

(A) Baryons (or heavy particles). Proton and particles heavier than proton form this group. Proton and neutron are called *nucleons* and the rest are called *hyperons*. Every baryon has an *antiparticle*. If a number, called the baryon number + 1 is assigned to baryons and a number – 1 is assigned to antibaryons, then in any closed system interaction or decay, the baryon number does not change. This is the law of conservation of baryons.

Hyperons is a special class of baryons characterised by a time decay of the order of 10^{-10} seconds and mass value intermediate between those of the neutron and deuteron. Their decay time is very much greater than the time of their formation (10^{-3} sec). It is because of this unsolved problem, that these particles, along with the *K*-mesons, are called *strange particles.* There are four types of hyperons; Lambda, Sigma, Xi and Omega.

TABLE 38.1. **Elementary particles**

	Name	*Symbol*	*Rest Energy M_0 / MeV*	*Mean lifetime τ/s*	*Common decay modes*
Leptons	Photon	γ	0	stable	
	Neutrino	ν_e	0	stable	
		ν_μ	0	stable	
	Electron	$e^\pm$	0.511004(2)	stable	
	Muon	$\mu^\pm$	105.659 (2)	$2.1994\ (6) \times 10^{-6}$	$e\,\nu\,\bar{\nu}$
Mesons	Pion	$\pi^\pm$	139.576 (11)	$2.602\ (2) \times 10^{-8}$	$\mu\nu$
		π^0	134.972 (12)	$0.84\ (10) \times 10^{-16}$	$\gamma\gamma$ (99%) $\gamma e^+ e^-$ (1%)
	Kaon	$K^\pm$	493.82 (11)	$1.235\ (4) \times 10^{-8}$	$\mu\,\nu$ (64%) $\pi^\pm \pi^0$ (21%) 3π (5%)
		K^0	497.76 (16)	50% K_1, 50% K_2	
		K_1		$8.62\ (6) \times 10^{-11}$	$\pi^+ \pi^-$ (69%) $2\pi^0$ (31%)
		K_2		$5.38\ (19) \times 10^{-8}$	$\pi e \gamma$ (39%) $\pi\mu\nu$ (27%) $3\pi^0$ (21%) $\pi^+ \pi^- \pi^0$ (13%)
	Eta	η^0	548.8 (6)		$\gamma\,\gamma$ (38%) $\pi\,\gamma\,\gamma$ (2%) $3\pi^0$ (31%) $\pi^+ \pi^- \pi^0$ (23%) $\pi^+ \pi^- \gamma$ (5%)
Baryons	Proton	$P^\pm$	938.256 (5)	Stable	
	Neutron	n	939.550 (5)	$9.32\ (14) \times 10^2$	pev
	Lambda	Λ^0	1115.60 (8)	$2.51\ (3) \times 10^{-10}$	$p\pi^2$ (65%) $n\pi^0$ (35%)
	Sigma	Σ^+	1189.4 (2)	$8.02\ (7) \times 10^{-11}$	$p\pi^0$ (52%) $n\pi^+$ (48%)
		Σ^0	1192.46 (12)	$< 10^{-14}$	$\Lambda\gamma$
		Σ^-	1197.32 (11)	1.49×10^{-10}	$n\pi^-$
	Xi	Ξ^0	1314.7 (7)	$3.03\ (18) \times 10^{-10}$	$\Lambda\pi^0$
		Ξ^-	1321.25 (18)	$1.66\ (4) \times 10^{-10}$	$\Lambda\pi^-$
	Omega	Ω^-	1672.5 (5)	$1.3\ (4) \times 10^{-10}$	$\Xi^0 \pi^-$, $\Xi^- \pi^0$, ΛK^- (?)

(B) Leptons. This group contains electron, photon, neutrino and muon.

(C) Mesons. The rest mass of these particles varies between about $250m_e$ and $1000m_e$. The mesons are the agents of interaction between particles inside the nucleus. Baryons and mesons, are jointly called *hadrons*, and are the particles of strong interaction.

38.2 Particles and Anti-particles

Electron and Positron. The positron and the electron are said to be antiparticles. They have the same mass and the same spin but opposite charge. They annihilate each other with the emission of photons, when they come in contact with each other. The existence of an antiparticle for the electron was actually predicted by Dirac, because of a symmetry of the equations of the relativistic quantum theory of the electron. Positron was discovered by Anderson in 1932.

Proton and antiproton. Dirac's theory, anticipating the positron, could be interpreted as implying a particle identical to the proton, except for a negative charge. The existence of this particle, the *antiproton*, was established in 1955 by Segre, Chamberlain, and their collaborators. Antiprotons were produced by bombarding protons in a target with 6-GeV protons, thereby inducing the reaction

$$p + p\,(+\text{energy}) \rightarrow p + p + p + \overline{p}$$

The K.E. of the bombarding proton is converted to a proton-antiproton pair plus the K.E. of the four residual particles. Antiprotons interact strongly with matter and annihilate with protons. In a typical annihilation reaction, the rest mass of the annihilating pair appears as five pions and their K.E.

$$p + \overline{p} \rightarrow \pi^+ + \pi^- + \pi^+ + \pi^- + \pi^0.$$

Neutron and antineutron. The antiparticle of neutron, *antineutron*, was discovered in 1956 by Cork, Lamberton and Wenzel. The nature of the antineutron is not very well known. Both neutron and antineutron have zero charge and the same mass. However, since neutron is supposed to have a certain internal charge distribution, it is expected that the antineutron has an internal charge distribution opposite to that of the neutron. Antineutron is quickly annihilated, either by a proton or a neutron, usually with the production of several pions. If an antineutron is not annihilated by a nucleon, it decays by the reaction

$$\overline{n} \rightarrow \overline{p} + \beta^+ + \nu\,.$$

Neutrino and antineutrino. As discussed in chapter 31, it is possible to avoid violations of the law of conservation of energy, linear and angular momenta and spin by assuming the existence of a particle, called a *neutrino*. It has *charge zero*, *spin* 1/2, *zero rest mass* (*rest mass very small in comparison with that of the electron*) *and magnetic moment smaller than* 10^{-8} *Bohr magneton or nearly zero.* The neutrino has a finite energy and momentum in flight. It travels with the speed of light c. It does not cause ionization on passing through matter.

The antiparticle of neutrino is *antineutrino.* The distinction between the neutrino ν and antineutrino $\bar{\nu}$ is a particularly interesting one. The spin of the neutrino is opposite in direction to the direction of its motion; viewed from behind, the neutrino spins-counterclockwise. But the spin of the antineutrino is in the same direction as its direction of motion; viewed from behind, it spins clockwise. Thus the neutrino moves through space in the manner of a left-handed screw, while the antineutrino does so in the manner of a right-handed screw. Thus a neutrino possesses a "left-handed" helicity; the antineutrino possesses a "right-handed" helicity, *i.e.*, *a neutrino and antineutrino differ only in the sense of their helicity.*

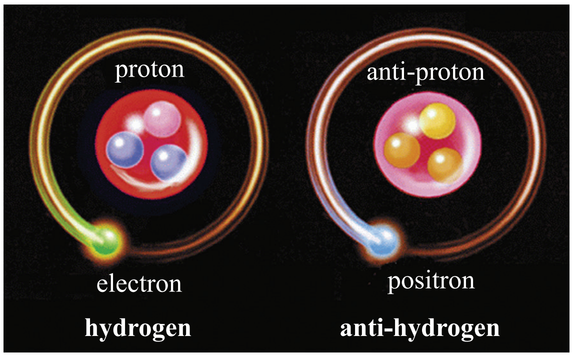
It is customary to call the particle accompanying a positron a neutrino, ν, while that accompanying an electron is called an antineutrino, $\bar{\nu}$. It is important to remember that positrons and electrons are never ejected alone. It is clear from experiments that neutrinos emitted in β decay (positron decay and K-capture) have a negative helicity, *i.e.*, they are longitudinally polarized with their spin axes antiparallel to their direction of travel.

$$\mathbf{H} = +1 \text{ for } \bar{\nu}\,;\ \mathbf{H} = - \text{ for } \nu\,.$$

Because of its lack of charge and magnetic moment, a neutrino has essentially no interaction with matter, except that which leads to inverse β decay. This interaction is extremely weak. The cross section for this process is only $\sigma \approx 10^{-48}\ \text{m}^2 \approx 10^{-20}$ *barn*. Matter is almost totally transparent to neutrinos.

38.3 Antimatter

In Atomic Physics, it has long been useful to consider an atom as composed of extranuclear electrons and a nucleus formed of protons and neutrons. There is reason to believe that a *positron* and an *antiproton* could form an atom of *antihydrogen*. Antihydrogen would have a spectrum similar to that of ordinary hydrogen. Indeed, from a collection of antiprotons, antineutrons, and positrons, a world of antimatter might be constructed, which could be indistinguishable from our world so long as everything were made of antiparticles. However, if some of this *contraterrene* matter were to come in contact with the ordinary matter, particle-antiparticle annihilation would occur with a tremendous release of energy.



38.4 The Fundamental Interactions

Four kinds of interaction between elementary particles can account for all known processes in the physical universe on all scales of size. In Table 38.2 the four basic interactions are summarised.

TABLE 38.2. The four fundamental interactions

Interaction	*Particles Affected*	*Range*	*Relative Strength*	*Particles Exchanged*
Strong	Hadrons	$\sim 10^{-15}$ m	1	Mesons
Electro-magnetic	Charged Particles	∞	$\sim 10^{-2}$	Photons
Weak	Hadrons and leptons	$\sim 10^{-17}$ m	$\sim 10^{-13}$	Intermediate bosons
Gravitational	All	∞	$\sim 10^{-40}$	Gravitons

(1) **Strong interaction.** A familiar example of strong interaction is the forces which hold nucleons together (nuclear forces) in the atomic nucleus. The strong nuclear interaction is independent of the electric charge. The range of these interactions is about 10^{-15} m. Time interval of such an interaction is roughly 10^{-23} s.

(2) **Electromagnetic interaction.** It operate on all charged particles. Thus electromagnetic interactions are charge dependent (attractive as well as repulsive). The range is infinite and the interaction works through the *photon*. The formation of electron-positron pair from γ-ray is an example of electromagnetic interaction.

(3) **Weak interaction.** All strong interactions take place in times of about 10^{-23} s (characteristic time). Yet is has been observed that some of the resulting particles, although energetically unstable, suffer no decay until a time 10^{13} times greater than 10^{-23} is reached. That is, their decay takes place in time of about 10^{-10} s. For example, β-decay of radioactive nuclei does not take place until a time 10^{13} times greater than that involved in strong interactions has approached. Had there been strong nuclear or electromagnetic interactions, there would have been no such delay in the decay process. Therefore, this delay in the decay process suggests that either these particles are not subjected to strong interacting forces or there is some new conservation law or prohibition which forbids the decay. But since most of the particles involved are subjected to either nuclear force or have electric charge or both, there must be some rule which stops the process. But eventually the decays do happen, there must be some other type of interaction as predicted by Fermi to explain β-decay. Since particles take long time to respond to such an interaction, force involved must be very weak

compared with strong nuclear forces. The range of such an interaction is less than 10^{-17} m. The characteristic time of this interaction is ≈ 10^{-10} s. Yukawa, in 1938, suggested that there should be a field quantum for the weak interaction, corresponding to the photon and pion. This so called *intermediate vector boson*. It has not been experimentally detected as a free particle. The weak interaction is responsible for the decay of strange and non-strange particles and for non-leptonic decays of strange particles.

(4) **Gravitational interaction.** It is the weakest of the four types of interactions. It has infinite range. Although gravity has a measurable influence on macroscopic bodies, its interaction with subnuclear particles is very small. Gravitation can be explained in terms of the interactions of '*gravitons*'. Their mass must be zero, and therefore, their velocity must be that of light. As the gravitational field is extremely weak, the gravitons can not be detected in the laboratory.

Of the four basic forces, only gravitational force is universal. Weak forces affect every particle other than photon. Electromagnetic force is confined to charged particles. Strong forces are the most selective and they serve as the criteria for classifying all known particles other than photons into two broad categories, the leptons and the hadrons. Leptons and photons are light particles and do not feel the strong forces. But hadrons feel the strong forces and participate in strong interactions besides taking part in other types of interactions as well. Thus the proton is a strongly interacting nuclear constituent and is therefore, a hadron. At the same time, because of its charge and mass, it must also experience electromagnetic and gravitational forces. The fact that it can be created by β-decay of neutron shows that the proton is involved in weak interactions as well.

38.5 Elementary-particle Quantum Numbers

In classifying the various elementary particles, several discrete quantum numbers are used. We are already familiar with two such quantum numbers, namely those that describe a particle's *charge* and *spin*. These quantum numbers specify measurable physical properties and are always conserved. We know that all elementary charges are 0 or ± 1. The charge is conserved in all processes and no exceptions are known. The spin quantum number J is either an integer or an half odd integer for the particles so far detected. Particles with integer spin obey the Bose-Einstein statistics and are called *bosons*. Particles with half odd integer spins obey the Fermi-Dirac statistics and are called *fermions*. The other quantum numbers are more abstract and it is not always clear precisely to what aspect of physical reality they refer.

(*a*) **Baryon number.** Each baryon is given a baryon number $B = 1$, each corresponding antibaryon is given a baryon number $B = -1$. All other particles have $B = 0$. The law of conservation of baryons states that the sum of the baryon numbers of all the particles after a reaction or decay must be the same as their sum before. This rule ensures that a proton cannot change into an electron, even though a neutron can change into a proton. Baryon conservation ensures the stability of the proton against decaying into a particle of smaller mass.

(*b*) **Lepton number.** Leptons are supposed to possess a property called *Lepton number* (L). Since the neutrinos associated with electrons and with muons are recognised as different, we introduce two lepton numbers L_e and L_μ, both of which must be conserved separately in particle reactions and decays. The number $L_e = 1$ is assigned to the electron and the e-neutrino and $L_e = -1$ to their antiparticles. All other particles have $L_e = 0$. Also, the number $L_\mu = 1$ is assigned to the muon and the μ-neutrino and $L_\mu = -1$ to their antiparticles. All other particles have $L_\mu = 0$. An example of particle-number conservation is the decay of the neutron, in which $B = 1$ and $L_e = 0$ before and after :

$$n \rightarrow p + e^- + \overline{\nu}_e$$

$$B : +1 \rightarrow (+1) + 0 + 0$$

$$L_e : 0 \rightarrow 0 + (+1) + (-1)$$

This is the only way in which the neutron can decay and still conserve both energy and baryon number B. Similarly, pion decay illustrates the conservation of L_μ :

$$\pi^- \rightarrow \mu^- + \bar{\nu}_\mu$$

$$L_\mu : 0 \rightarrow (+1) + (-1)$$

(*c*) **Strangeness number.** The properties of K-mesons and hyperons appeared at first instance to be paradoxical in that they were produced by high energy reactions but always in pairs *i.e.*, if one particle of some kind is produced then simultaneously another different particle is also emitted. This is termed as the *associated production* of kaons and hyperons which are known as *strange particles*. Furthermore as they are produced in strong interaction, one should expect them to decay also through strong interactions in a very short time but this is not observed. Instead they decay slowly. Because of this strange behaviour they were called as strange particles. Their production is typical of strong interaction such as :

$$\pi^- + p^+ \rightarrow \wedge^0 + K^0$$

and their decay

$$\wedge^0 \rightarrow p + \pi^-$$

$$K^0 \rightarrow \pi^+ + \pi^-$$

is through weak interaction.

This and other considerations led to the notion of *strangeness number* S. S = 0 for the nucleons, and S ≠ 0 for kaons and hyperons (Table 38.3). It is found that S is conserved in all processes mediated by the strong and electromagnetic interactions. Kaons and hyperons, which have S ≠ 0, are created in high-energy collisions that involve the strong interaction. Their multiple appearance results from the necessity to conserve S. An example is the proton-proton collision :

$$p^+ + p^+ \rightarrow \wedge^0 + K^0 + p^+ + \pi^+$$

$$S : 0 + 0 \rightarrow (-1) + 1 + 0 + 0$$

On the other hand, S can change in an event governed by the weak interaction. The decays of kaons and hyperons proceed via the weak interaction and accordingly are extremely slow. Even the weak interaction, however, is unable to permit S to change by more than ± 1 in a decay.

(*d*) **Hypercharge.** A quantity called *hypercharge* (Y) is conserved in strong interaction. Hypercharge is equal to the sum of the strangeness and baryon numbers of the particle families : $Y = S + B$. The various hypercharge assignments are listed in Table 38.3.

TABLE 38.3

Class	*Name*	*Symbol*	*Spin*	*B*	*Le*	L_μ	*S*	*Y*	*I*
LEPTON	*e*-neutrino	ν_e	$\frac{1}{2}$	0	+ 1	0			
	µ-neutrino	ν_μ	$\frac{1}{2}$	0	0	+ 1			
	Electron	e^-	$\frac{1}{2}$	0	+ 1	0			
	Muon	μ^-	$\frac{1}{2}$	0	0	+ 1			
MESON	Pion	$\pi^\mp$ π^0	0	0	0	0	0	0	1
	Kaon	K^+ K^o	0	0	0	0	+ 1	+ 1	$\frac{1}{2}$
	η meson	η^0	0	0	0	0	0	0	0

BARYON	Nucleon (Proton) Neutron)	p n	$\frac{1}{2}$	+ 1	0	0	0	+ 1	$\frac{1}{2}$
	Λ hyperon	Λ^0	$\frac{1}{2}$	+ 1	0	0	– 1	0	0
	Σ hyperon	Σ^+ Σ^0 Σ^-	$\frac{1}{2}$	+ 1	0	0	– 1	0	1
	Ξ hyperon	Ξ^0 Ξ^-	$\frac{1}{2}$	+ 1	0	0	– 2	– 1	$\frac{1}{2}$
	Ω hyperon	Ω^-	$\frac{3}{2}$	+ 1	0	0	– 3	– 2	0

(*e*) **Isospin and Isospin quantum number.** As far as strong interactions are concerned, the neutron and the proton are two states of equal mass of a nucleon doublet. From Table 38.1, it is found that particles occur in multiplets. For example, singlets η^0, Ω^-, $\wedge^0$. Doublet : p, n, triplet π^0, π^-, π^+. It is natural to think of the members of a multiplet as representing different charge states of a single fundamental entity. It has proved useful to categorize each multiplet according to the number of charge states it exhibits by a number I such that the multiplicity of the state is given by $2I + 1$. Thus the nucleon multiplet is assigned $I = 1/2$, and its $2 \times 1/2 + 1 = 2$ states are the neutron and the proton. The pion multiplet has $I = 1$, and its $2 \times 1 + 1 = 3$ states are π^+, π^0 and π^-.

Isospin can be represented by a vector I in "isospin space" whose component in any specified direction is governed by a quantum number customarily denoted I_3. The possible values of I_3 are restricted to $I, I - 1, I - 2$.........., 0, ...– $(I - 1), - I$. Hence I_3 is half-integral, if I is half-integral and integral or zero if I is integral. For the nucleon, $I = 1/2$ and $I_3 = 1/2$ corresponds to the proton, $I_3 = - 1/2$ for the neutron. Similarly, for the pion triplet $1 = I$ giving $I_3 = + 1$ for π^+, 0 for π and – 1 for π^-. The charge of a meson or baryon is related to its baryon number B, its strangeness number S, and the component I_3 of its isospin by the formula

$$q = e\left(I_3+\frac{B+S}{2}\right)= e\left(I_3+\frac{Y}{2}\right)$$

38.6 Conservation Laws and Symmetry

A very important set of conservation laws is related to symmetries involving parity (P), charge conjugation (C), and time reversal (T).

Conservation of parity. Parity relates to the symmetry of the wave function that represents the system. If the wave function is unchanged, when the coordinates (x, y, z) are replaced by $(- x, - y, - z)$, then the system has a parity of + 1. If the wave function has its sign changed, when the coordinates are reversed, then the system has a parity of – 1. If we write $\psi (x, y, z) = P\psi (- x, - y, - z)$, we can regard P as a quantum number characterizing ψ whose possible values are + 1 and – 1. During a reaction in which parity is conserved, the total parity number does not change.

Changing the coordinates (x, y, z) into $(- x, - y, - z)$ converts a right-handed coordinate system into a left-handed coordinate system. In terms of symmetry, the meaning of conservation of parity is that in any situation where parity is conserved, the description of the reaction will not be changed if the word "left" is changed to the word "right" and vice versa. This means that such reactions can provide no clue that will distinguish between the directions right and left. Prior to 1956 it was believed that all reactions in nature obeyed the law of conservation of parity. However, Yang and

Lee pointed out that in reactions involving the weak interaction, parity was not conserved, and that experiments could be devised that would absolutely distinguish between right and left. Indeed parity conservation is found to hold true only in the strong and electromagnetic interactions.

Charge conjugation symmetry. Charge conjugation is the act of symmetry operation in which every particle in a system is replaced by its antiparticle. If the anti-system, or antimatter counterpart exhibits the same physical phenomena, then charge parity (C) is conserved. For example, if in a hydrogen atom, the proton is replaced by an antiproton and the electron is replaced by a positron, then this antimatter atom will behave exactly like an ordinary atom, if observed by people also made of antimatter. In fact C is not conserved in the weak interaction.

Time reversal symmetry. Time parity T describes the behaviour of a wave function when t is replaced by $-t$. The symmetry operation that corresponds to the conservation of time parity is time reversal. Time reversal symmetry implies that the direction of time is not significant, so that the reverse of any process that can occur is also a process that can occur. In other words, if symmetry under time reversal holds, it is impossible to establish by viewing it whether a motion picture of an event is being run forward or backward. Prior to 1964, time parity T was considered to be conserved in every interaction. It was discovered in 1964 that one form of the K^0 kaon can decay into $\pi^+ + \pi^-$, which violates the conservation of T. The symmetry of phenomena under time reversal thus does not seem to be universal.

Combined inversion of CPT. The combined symmetry operation in which the antimatter mirror-image of a system is run in reverse allows a test of CPT invariance. All the evidence supports the conservation of CPT. The conservation of CPT means that for every process there is an antimatter mirror-image counterpart that takes place in reverse. This particular symmetry seems to hold for all interactions, even though its component symmetries sometimes fail individually.

38.7 The Quark Model

Murray Gell-Mann and G. Zweig proposed the quark model in 1964. This theory is based on the idea that the hadrons are built up from a limited number of "fundamental" units, which have acquired the name quarks. The original three quarks were labeled u (for "*up*"), d (for "*down*") and s (for "*strange*").

u quark has electric charge $+\frac{2}{3}e$ and strangeness 0.

d quark has electric charge $-\frac{1}{3}e$ and strangeness 0.

s quark has electric charge $-\frac{1}{3}e$ and strangeness -1.

Each quark has a baryon number of $B = 1/3$.

Each quark has an antiquark associated with it $(\bar{u}, \bar{d}, \text{and } \bar{s})$. The magnitude of each of the quantum numbers for the antiquarks has the same magnitude as those for the quarks, but the sign is changed.

Compositions of hadrons according to the quark model

Hadrons may be *baryons* or *mesons*. A baryon is made up of three quarks. For example, the proton is made up of two u quarks and a d quark (uud). For these quarks, the electric charges are $+2/3$, $+2/3$, and $-1/3$, for a total value of $+1$. The baryon numbers are $+1/3$, $+1/3$ and $+1/3$, for a total of $+1$. The strangeness numbers are 0, 0 and 0 for a total strangeness of 0. All are in agreement with the quantum numbers for the proton. Fig. 38.1 shows quark models of the proton, antiproton, neutron and antineutron. Electric charges are given in units of e.

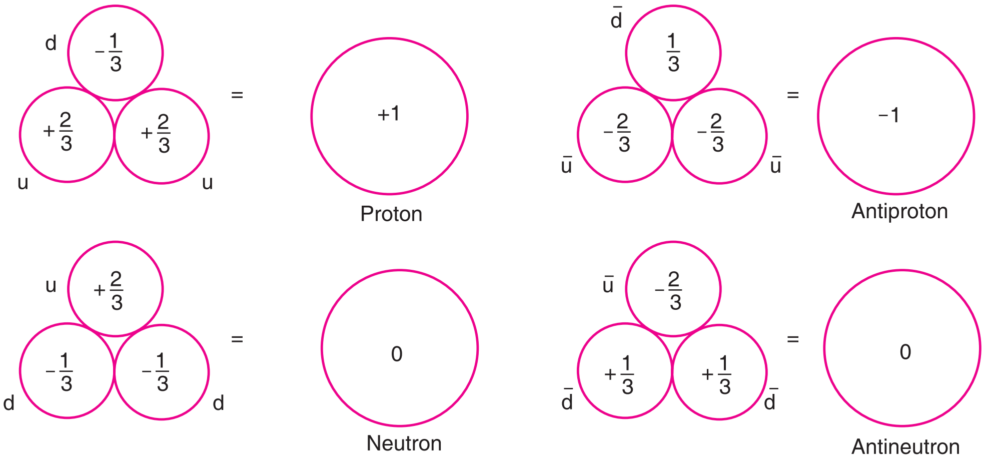
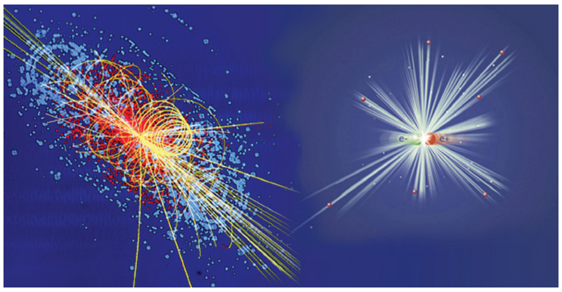


Fig. 38.1

A meson is made up of one quark and one antiquark. For example, the π^+ **meson** is the combination of a u quark and a d antiquark $(u\bar{d})$. Electric charges of these quarks are + 2/3 and + 1/3 for a total of + 1. The baryon numbers are + 1/3 and – 1/3 for a total baryon number of 0. The strangeness numbers are 0 and 0 for a total of 0. All of these are in agreement with the quantum numbers for the pi-meson. Quarks all have spins of 1/2, which accounts for the observed half-integral spins of baryons and the 0 or 1 spins of mesons.



Each Collision between Hadrons is actually a group of Quark Collisions with uncertain Energies.

All known hadrons can be explained in terms of the various quarks and their antiquarks. Table 38.4 shows the quark contents of five hadrons and how they account for the observed charges, spins, and strangeness numbers of these particles.

TABLE 38.4. Compositions of some hadrons according to the Quark Model

Hadron	*Quark content*	*Baryon number*	*Charge, e*	*Spin*	*Strangeness*
π^+	$u\bar{d}$	$\frac{1}{3}-\frac{1}{3}=0$	$+\frac{2}{3}+\frac{1}{3}=+1$	$\uparrow\downarrow=0$	$0+0=0$
K^+	$u\bar{s}$	$\frac{1}{3}-\frac{1}{3}=0$	$+\frac{2}{3}+\frac{1}{3}=+1$	$\uparrow\downarrow=0$	$0+1=+1$
P^+	uud	$\frac{1}{3}+\frac{1}{3}+\frac{1}{3}=+1$	$+\frac{2}{3}+\frac{2}{3}-\frac{1}{3}=+1$	$\uparrow\uparrow\downarrow=\frac{1}{2}$	$0+0+0=0$
n^0	ddu	$\frac{1}{3}+\frac{1}{3}+\frac{1}{3}=+1$	$-\frac{1}{3}-\frac{1}{3}+\frac{2}{3}=0$	$\downarrow\downarrow\uparrow=\frac{1}{2}$	$0+0+0=0$
Ω^-	sss	$\frac{1}{3}+\frac{1}{3}+\frac{1}{3}=+1$	$-\frac{1}{3}-\frac{1}{3}-\frac{1}{3}=-1$	$\uparrow\uparrow\uparrow=\frac{3}{2}$	$-1-1-1=-3$

Coloured quarks and gluons : There were problems with the quark model, one of them being Ω^- hyperon. It was believed to contain three identical s quarks (sss). This violates the Pauli exclusion principle, that prohibits two or more fermions from occupying identical quantum states. The proton, neutron, and others with two identical quarks would violate this principle also. We can resolve this difficulty by assigning a new property to the quarks. We can regard this new property as an additional quantum number that can be used to label the three otherwise identical quarks in the Ω^-. If this additional quantum number can take any one of three possible values, we can restore the Pauli principle by giving each quark a different value of this new quantum number, which is known as *colour*. The three colours are labeled red (R), blue (B), and green (G). The Ω^- for example, would then $s_R\, s_B\, s_G$. The antiquark colours are antired (R) antiblue (B) and antigreen (G).

An essential component of the quark model with colour is that *all observed meson and baryon states are "colourless"*, *i.e.*, either colour anticolour combinations in the case of mesons, or equal mixtures of R, B and G in the case of baryons.

Since hadrons seem to be composed of quarks, the strong interaction between hadrons should ultimately be traceable to an interaction between quarks. The force between quarks can be modeled as an exchange force, mediated by the exchange of massless spin – 1 particles called gluons. Eight gluons have been postulated. The field that binds the quarks is a *colour field. Colour is to the strong interaction between quarks as electric charge is to the electromagnetic interaction between electrons.* It is the fundamental strong "charge" and is carried by the gluons. The gluons must therefore be represented as combinations of a colour and a possibly different anticolour. The gluons are massless and carry their colour-anticolour properties just as other particles may carry electric charge. For example, Fig. 38.2 shows a gluon $R\bar{B}$ being exchanged by red and blue quarks. In effect the red quark emits its redness into a gluon and acquires blueness by also emitting antiblueness. The blue quark, on the other hand, absorbs the $R\bar{B}$ gluon, canceling its blueness and acquiring a red colour in the process.

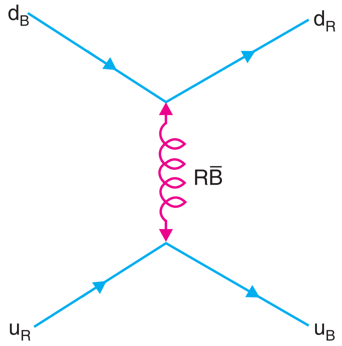
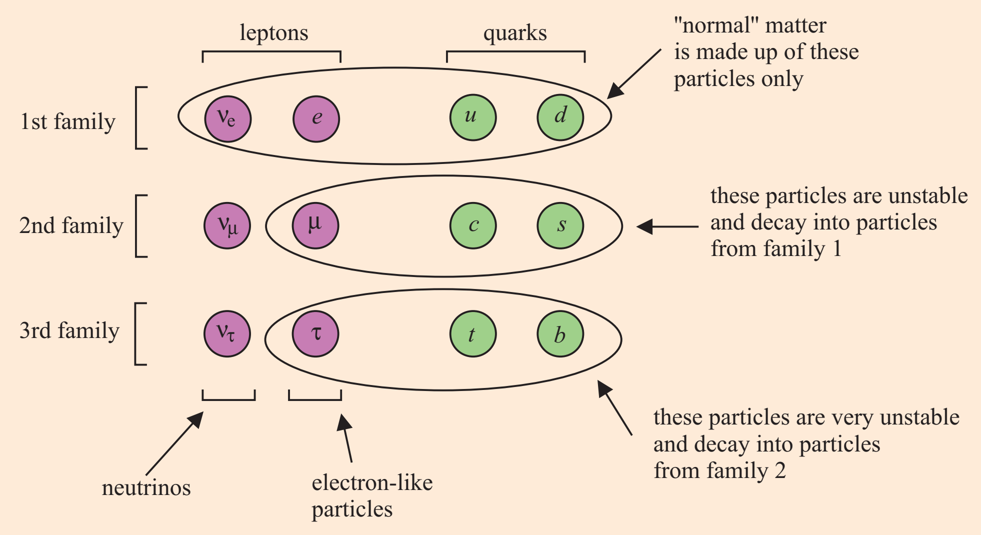

d_B
d_R
$R\bar{B}$
u_R
u_B

Fig. 38.2

Charm, Bottom, and Top. In 1970, Glashow, Iliopoulis, and Maiani proposed the existence of a fourth quark, called c or *charmed* quark. The charmed quark was suggested to explain the suppression of certain decay processes that are not observed. With only three quarks, the processes would proceed at measurable rates and should have been observed. The charm quark has a charge of $\frac{2}{3}e$, strangeness 0 and a charm quantum number of + 1. Other quarks have 0 charm.


leptons
quarks
"normal" matter
is made up of these
particles only
1st family
ν_e
e
u
d
2nd family
ν_μ
μ
c
s
these particles are unstable
and decay into particles
from family 1
3rd family
ν_τ
τ
t
b
these particles are very unstable
and decay into particles
from family 2
neutrinos
electron-like
particles

In 1977, a new particle was discovered at Fermi Lab that provided evidence for yet another quark. This particle, called the upsilon-meson, was thought to be made up of the new quark called b (for bottom or beauty) along with the associated antiquark $\bar{b}$. b quark has electric charge –1/3e.

Because quarks seem to come in pairs, it is expected that there is a partner to the b quark, called t (for top, if b = bottom, or truth, if b = beauty). It has a charge of $+2/3e$.

Three generations of quarks and leptons. Both leptons and quarks appear to come in three generations of doublets, with all particles having spin 1/2. Table 38.5 shows the properties of the three generations of quarks and leptons. The first generation contains two leptons, the electron and the electron neutrino, and two quarks, up and down. All the properties of ordinary matter can be understood on the basis of these particles. The second generation includes the muon and muon-neutrino and the charm and strange quarks. These particles are responsible for most of the unstable particles and resonances created in high energy collisions. The third generation includes the tau and the tau-neutrino and the top and bottom quarks.

TABLE 38.5. Properties of the three generations of Quarks and Leptons

Generation	*Quark*	*Symbol*	*Charge, e*	*Strangeness*	*Charm*
1	Up	u	$+\frac{2}{3}$	0	0
	Down	d	$-\frac{1}{3}$	0	0
2	Charm	c	$+\frac{2}{3}$	0	+ 1
	Strange	s	$-\frac{1}{3}$	– 1	0
3	Top	t	$+\frac{2}{3}$	0	0
	Bottom	b	$-\frac{1}{3}$	0	0

Generation	*Lepton*	*Symbol*	*Charge, e*
1	Electron	e^-	– 1
	e-Neutrino	ν_e	0
2	Muon	μ^-	– 1
	μ-Neutrino	ν_μ	0
3	Tau	τ^-	– 1
	τ-Neutrino	ν_τ	0

EXERCISE

1. Write short notes on (*a*) Elementary particles (*b*) Conservation laws in elementary particles (*c*) Weak nuclear interactions (*d*) Strange particles.

2. Which of the following reactions can occur? State the conservation laws violated by the others.

(*a*) $p + p \rightarrow n + p + \pi^+$ (*b*) $p + p \rightarrow p + \Lambda^0 + \Sigma^+$ (*c*) $e^+ + e^+ \rightarrow \mu^+ + \pi^-$

(*d*) $\Lambda^0 \rightarrow \pi^+ + \pi^-$ (*e*) $\pi^- + p \rightarrow n + \pi^0$

[**Ans.** (*a*) and (*e*) can occur; (b) does not conserve B or spin; (*c*) does not conserve L_e, L_μ or spin; (*d*) does not conserve B or spin.]

3. If the pion decays from rest to give a muon of 4.05 MeV energy, what is the K.E. of the accompanying neutrino? What is the mass of the neutrino in this process? [**Ans.** 30.12 MeV : 59 m_e]

4. When an antiproton of energy 72 MeV comes to rest in a nuclear emulsion it is annihilated by a proton to form two pairs of positive and negative pions and a neutral pion. Calculate the average K.E. of each pion, assuming they all have the same energy.

[**Hint.** The surplus energy available to the reaction products = (938.3 MeV + 72 MeV + 938.3 MeV) – (2 × 139.6 + 2 × 139.6 + 135) = 1255 MeV. This energy is distributed among the 5 particles and on an average each shares ≈ 250 MeV.]

5. Using the law of conservation of lepton numbers, find which of the following reactions is possible:

(*a*) $p+\overline{v}_e \rightarrow n+\mu^+$ (*b*) $p+\overline{v}_e \rightarrow n+e^+$ [**Ans.** *b* is possible]

6. Using the baryon number and the strangeness number conservation laws, find which of the following reactions is allowed :

(*a*) $\pi^- + p \rightarrow \wedge^0 + K^0$ (*b*) $\pi^- + p \rightarrow \wedge^0 + \pi^0$ [**Ans.** *a* is allowed]

7. Find the maximum kinetic energy of the electron emitted in the beta decay of the free neutron. The neutron-proton mass differences is 1.30 MeV. [**Ans.** $n \rightarrow p^+ + e^- + \overline{v}_e + Q$

$$Q = M_n - M_p - m_e = (1.30 \text{ MeV}) - (0.51 \text{ MeV}) = 0.79 \text{ MeV}.$$

This energy-release goes into the kinetic energy of the electron plus the kinetic energy of the antineutrino. The two particles can share the 0.79 MeV of energy in any way they choose. Thus the electron may have a maximum kinetic energy of **0.79 MeV.**]

8. A μ^- meson collides with a proton, and a neutron plus another particle are created. What is the other particle? [**Ans.** $\mu^- + p^+ \rightarrow n^0 + ?$

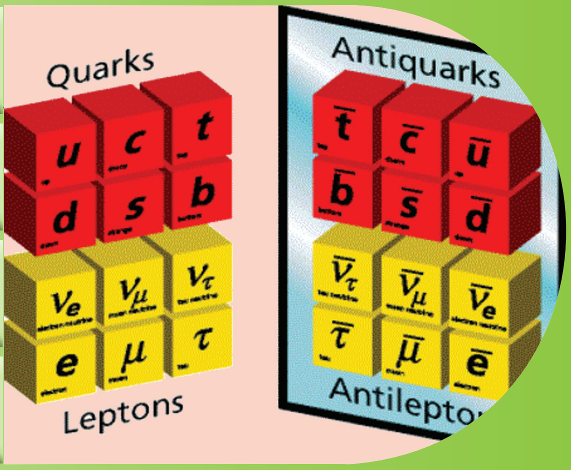
The other particle must have zero charge, spin $\frac{1}{2}$ and $L = 1$. Hence it can only be a neutrino.]

9. Why must the quarks in a hadron have different colours? Would they have to have different colours if their spins were 0 or 1 rather than $\frac{1}{2}$? [**Ans.** In order to obey the exclusion principle; no.]

10. What would be the structure of K$^+$ and Σ$^+$ in terms of the quark model? The strangeness quantum number of each of the quarks *u* and *d* is zero and that of the quark *s* is – 1. [**Ans.** $\text{K}^+ (= \overline{\text{s}}\text{u}), \Sigma^+ (= \text{suu})$]

CHAPTER

39


Quarks
u c t
d s b
ν_e ν_μ ν_τ
e μ τ
Leptons
Antiquarks
$\bar{t}$ $\bar{c}$ $\bar{u}$
$\bar{b}$ $\bar{s}$ $\bar{d}$
$\bar{\nu}_\tau$ $\bar{\nu}_\mu$ $\bar{\nu}_e$
$\bar{\tau}$ $\bar{\mu}$ $\bar{e}$
Antilepto

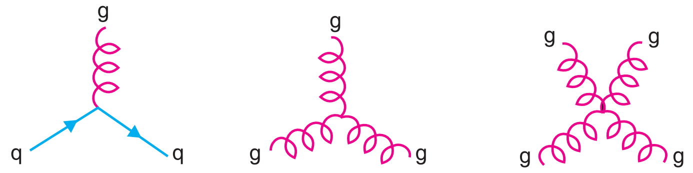
ELEMENTARY PARTICLES–II

AT A GLANCE

39.1 Quantum Chromodynamics
39.2 The Standard Model
39.3 Unification of Interactions

39.1 Quantum Chromodynamics

The theory of the quark gluon interaction is called *Quantum Chromodynamics* (QCD). Let us draw the Feynman diagrams for the quark-gluon interaction. Quarks are spin $-\frac{1}{2}$ ferminons and, gluons are spin-1 bosons. The basic vertex consists of the meeting of two fermion arms and a boson arm (Fig. 39.1(*a*)). But there are two other vertices, one with three gluon arms (Fig. 39.1(*b*)) and a second with four gluon arms (Fig. 39.1(*c*)). Thus a gluon can interact directly with another gluon. Thus 3- or 4-gluon vertices exist in QCD. Their appearance in QCD is a consequence of the non-Abelian nature of gauge invariance in QCD.


g
q
q
g
g
g
g
g
g
g

(*a*) (*b*) (*c*)

Fig. 39.1

The quark-quark interaction is now by gluon exchange as in Fig. 39.2.

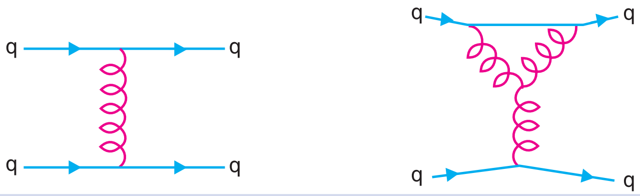


Fig. 39.2

It is believed that another consequence of the structure of QCD is that the central part of the potential between two quarks increases indefinitely as the distance between them increases, so that it is impossible to separate bound quarks: quarks are said to be *confined.* The fact that no free quark has been observed even in the most energetic collisions observable is supposed to be due to quark confinement.

QCD seems able to explain the behaviour of quarks within hadrons and has predicted certain effects that have been observed in high-energy experiments. But it has not yet been successful in accounting in a detailed way for that part of the interaction between quarks that leads to hadron-hadron forces.

39.2 The Standard Model

The *standard model* is the name given to the present model that we have of the elementary particles and their interactions. In this model, the theory of the strong interaction has been joined to that of the electroweak interaction to make a single picture. In this model the strong, weak, and electromagnetic forces appear as different manifestations of one basic phenomenon, with leptons and quarks finding natural places within the scheme. By being able to include leptons and quarks in the same framework, it becomes possible to explain, why the electron (a lepton) and the proton (a composite of quarks) have electric charges of exactly the same magnitude.

The ingredients of this model are given in Table 39.1.

Table 39.1. The particles and interactions of the standard model

Particles

1. Spin $-\frac{1}{2}$ fermions and antifermions :
    - (*a*) three generations of leptons (e^-, ν_e), (μ^-, ν_μ) and (τ^-, ν_τ),
    - (*b*) three generations of quarks (d, u), (s, c) and (b, t),

plus all their antiparticles.

2. Spin 1 gauge bosons :
    - (*a*) one massless electroweak boson, the photon γ ,
    - (*b*) three massive electroweak bosons, W^+, W^-, and Z^0,
    - (*c*) eight coloured gluons.
3. Spin-0 Higgs bosons.

Interactions

1. The electromagnetic coupling of the photon.
2. The interactions of the bosons, $W^\pm$, Z^0.
3. The strong interactions of the gluons with gluons and with the quarks.

The electroweak theory unifies 1 and 2.

(*i*) We have discussed leptons and quarks in chapter 38.

(*ii*) Gauge Bosons. These particles are the carriers of the interaction between the fundamental fermions and are the modern version of Yukawa's boson as carrier of a force. We know twelve guage bosons : 8 gluons, the photon, the W^+, W^-, and the Z^0.

The gauge bosons are the particles responsible for the interactions between fermions : the gluons are coupled only to the quarks, whereas the massive electroweak bosons (Z^0, $W^\pm$) are coupled to all the fermions. The massless electroweak boson (γ) is coupled to all electrically charged particles.

(*iii*) The Higgs Boson. It is the spin-0 boson which must exist if the Higgs mechanism is responsible for giving mass to the gauge bosons $W^\pm$, Z^0. The *Higgs mechanism* is the mechanism by which the existence of a spin-0 particle can give a gauge boson mass without breaking the gauge symmetry. Higgs showed that a field, called the *Higgs field*, must exist everywhere in space. By interacting with the Higgs field, particles acquire their characteristic masses. The stronger the interaction, the greater the mass.

The Higgs boson mediates the action of the Higgs field. The mass of the Higgs boson could be 1 TeV/c^2.

39.3 Unification of Interactions

The following four types of fundamental interactions are known at present :

1. *Gravitational interaction* 2. *Electromagnetic interaction* 3. *Strong interaction* 4. *Weak interaction.*

There have been efforts to unify the various interactions and deduce them as different manifestations of the same underlying theory. Such a unification was achieved for electric and magnetic forces by Maxwell.

Similarly, the physicists have succeeded in unifying the weak and electromagnetic interactions. The electromagnetic interaction, say, between two electrons, is due to the exchange of photons. Likewise, the weak interaction, for example, between an electron and a neutrino and between a proton and a neutrino, occurs through the exchange of one of a set of particles, called W$^+$, W$^-$ and Z^0 particles. For instance, in the β-decay of a neutron, a virtual W$^-$ is emitted, that is,

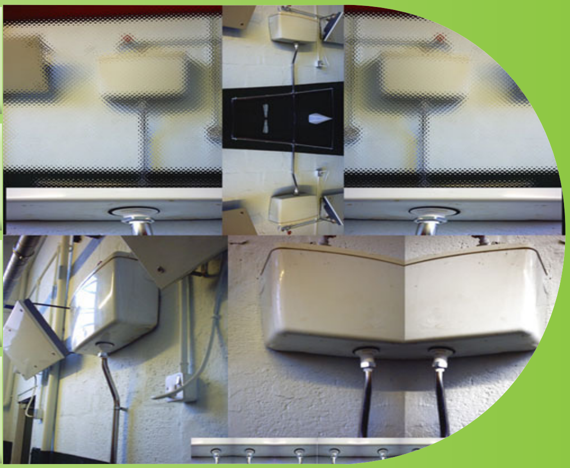
The Gravitational Interaction between Galaxies can Drastically change their Global Properties.

$$n \rightarrow p + W^- \text{ (virtual)} \rightarrow p + e^- + \bar{\nu}_e .$$

Since the weak interaction has a short range, each of the particles W^+, W^- and Z^0 is massive. The masses of W^+, and W^- are about 80,000 MeV/c^2 and the mass of Z^0 is about 90,000 MeV/c^2. The existence of these particles was predicted by A. Salam and S. Weinberg and was subsequently confirmed during the proton-antiproton collisions conducted at CERN.

Theoretically, the strong interaction has been unified with the weak electromagnetic interactions. This has been done by ensuring that the strong interaction is mediated by a set of particles, called gluons. However, the presence of gluons has not yet been experimentally confirmed though there is a strong evidence that such particles exist.

Finally there is the gravitational interaction which is generally insignificant for interactions between elementary particles.



CHAPTER 40

INTERACTION OF CHARGED PARTICLES WITH MATTER

AT A GLANCE

40.1 Rate of Loss of Energy of a Charged Particle Trasversing a Material Medium
40.2 Bremsstrahlung
40.3 Neutron Decay
40.4 Stability of Proton
40.5 Resonance Particles
40.6 Detection of Neutrino

40.1 Rate of Loss of Energy of a Charged Particle Trasversing a Material Medium

Let a particle of charge Ze move past a stationary electron with a velocity v (Fig. 40.1). Let v be so high that the electron may be assumed to remain stationary as the particle passes by. Let the distance of closest approach of the particle to the electron be b. Let m be the mass of the electron.

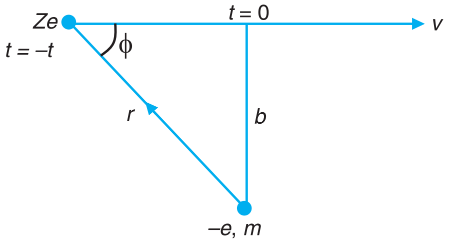


Fig. 40.1

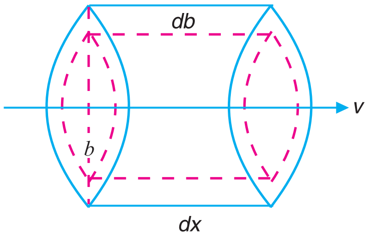


Fig. 40.2

The coulomb force between the particle and the electron $F = \dfrac{Ze^2}{4\pi\varepsilon_0 r^2}$.

If this force is exerted on the electron over a small interval of time dt, then the impulse ($F\,dt$) imparted to it is equal to the momentum (p) gained by the electron.

Resolving F into $F \cos \phi$ parallel to v and $F \sin \phi$ perpendicular to v, we find from symmetry consideration that the cosine components cancel out. p is, therefore, due to the component of F transverse to v. Hence,

$$p = \int_{-\infty}^{+\infty} (F \sin \phi)\, dt = \int_{-\infty}^{+\infty} \left(\frac{Ze^2}{4\pi\varepsilon_0 r^2} \sin \phi \right) dt \qquad \text{...(1)}$$

From Fig. 40.1, $\quad -vt = b \cot \phi$

$$\therefore \quad dt = \frac{b \operatorname{cosec}^2 \phi}{v} \cdot d\phi$$

From the Fig., $\quad \dfrac{1}{r} = \dfrac{\sin \phi}{b} \qquad \therefore \quad \dfrac{1}{r^2} = \dfrac{\sin^2 \phi}{b^2}$

The limits of integration are from 0 to π for the ϕ variable. We treat v as a constant.

Substituting these values in Eq. (1), we get

$$p = \frac{Ze^2}{4\pi\varepsilon_0 bv} \int_0^{\pi} \sin \phi \, d\phi = \frac{Ze^2}{2\pi\varepsilon_0 bv}$$

The gain of kinetic energy of the electron is

$$E = \frac{p^2}{2m} = \frac{Z^2 e^4}{8\pi^2 \varepsilon_0{}^2 m b^2 v^2} \qquad \text{...(2)}$$

Let n be the number of atoms in unit volume of the medium. Let Z' be the atomic number of medium. Consider a cylindrical shell in the material. Let b and $b + db$ be the inner and outer radii of the shell. Let dx be its length (Fig. 40.2).

Volume of the shell = $2\pi b\, db\, dx$

The number of electrons in this shell $= (2\pi b\, db\, dx)\, nZ'$ ...(3)

The kinetic energy (dE) gained by all electrons in the shell is

$$dE = \frac{Z^2 e^4}{8\pi^2 \varepsilon_0^2 m b^2 v^2} \times (2\pi b\, db\, dx\, nZ') = \frac{Z^2 e^4 nZ' dx}{4\pi \varepsilon_0^2 m v^2} \left(\frac{db}{b} \right)$$

The gain of kinetic energy of electrons equals the loss of energy of the heavy charged particle passing through the medium.

Thus, the energy lost by the charged particle in traversing unit distance in the medium is

$$\frac{-dE}{dx} = \frac{Z^2 e^4 nZ'}{4\pi \varepsilon_0^2 m v^2} \int_{b'}^{b''} \left(\frac{db}{b} \right) = \frac{Z^2 e^4 nZ'}{4\pi \varepsilon_0^2 m v^2} \ln \frac{b''}{b'} \qquad \text{...(4)}$$

b' and b'' are the lower and the upper limits respectively of the impact parameter b.

To find b′ : In a head on elastic collision it follows from the laws of conservation of energy and momentum that the speed imparted to the electron by the heavy charged particle is $2v$.

The kinetic energy of the electron $= \dfrac{1}{2} m (2v)^2 = 2mv^2$

Equating this to Eq. (2) with $b = b'$,

$$2mv^2 = \frac{Z^2 e^4}{8\pi^2 \varepsilon_0^2 m b'^2 v^2}$$

$$\therefore \qquad b' = \frac{Ze^2}{4\pi\varepsilon_0 m v^2} \qquad \text{...(5)}$$

To find b'' : b'' corresponds to the maximum impact parameter when the electrons' kinetic energy is equal to the average value of excitation (and ionisation) energies of the atom in the medium. We denote this average ionisation energy as I. Thus, using Eq. (2) with $b = b''$,

$$I = \frac{Z^2 e^4}{8\pi^2 \varepsilon_0^2 m b''^2 v^2}$$

$$b'' = \frac{Ze^2}{8\pi\varepsilon_0 v}\left(\frac{2}{mI}\right)^{1/2} \qquad \text{...(6)}$$

$$\frac{b''}{b'} = \left(\frac{2mv^2}{I}\right)^{1/2}$$

Eq. (4) becomes

$$-\frac{dE}{dx} = \frac{Z^2 e^4 nZ'}{4\pi\varepsilon_0^2 m v^2} \ln\left(\frac{2mv^2}{I}\right)^{1/2} \qquad \text{...(7)}$$

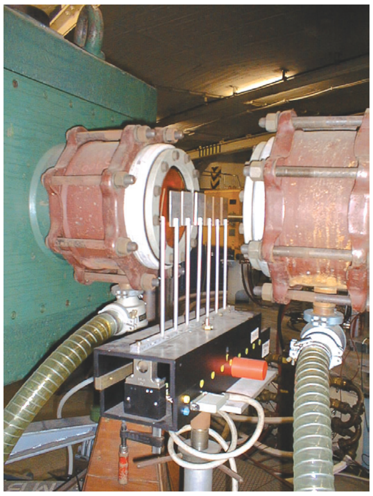
Eq. (7) gives the rate of loss of energy of the charged particle traversing a material medium.

Range. The total range of a particle in a given material is the length of the flight path before the particle comes to rest. It is given by the relation

$$R = \int_E^0 dx = \int_E^0 \left(-\frac{dE}{dx}\right)^{-1} (-dE) = \int_0^E -\left(\frac{dE}{dx}\right)^{-1} dE.$$

40.2 Bremsstrahlung

From electromagnetic theory, it is known that acceleration (or deceleration) of charged particles produces electromagnetic radiation. When a charged particle collides with the atom, due to interaction between the particle and the nucleus of the atom, it decelerates, thereby producing radiation (Fig. 40.3). This phenomenon in which radiation is produced due to deceleration of charged particle in a medium is called *Bremsstrahlung*. The loss of energy due to such radiation (*Bremsstrahlung*) becomes important only when (*a*) we take absorbers of high atomic number, (*b*) the energy of the incident particle is high, and (*c*) the mass of the incident particle is small.



Bremsstrahlung from an Internal Target Produces the Electrons.

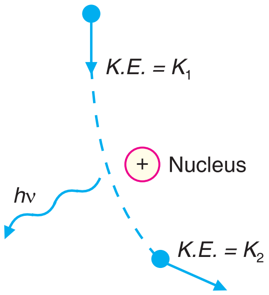


Fig. 40.3

For example, production of X-rays from a target material (on which fast moving electrons collide) is an example of Bremsstrahlung. The electrons are slowed down as they pass close to the nucleus of the atom and they radiate energy.

40.3 Neutron Decay

A free neutron is not stable and it decays into a proton, an electron and an anti-neutrino with a half life of 11 minutes.

The reaction, $n \rightarrow p + e^{-} + \overline{\nu}_e$

is allowed by both the conservation law of lepton and baryon numbers.

$$n \rightarrow p + e^{-} + \overline{\nu}_e$$
$$L = 0 \; 0 \; 1 \quad -1 \; (L = 0)$$
$$B = 1 \; 1 \; 0 \quad\;\; 0 \; (B = 1)$$

In this process, mass energy is conserved. Electric charge is conserved due to the presence of p and e^{-}. Angular momentum is conserved due to the presence of $\overline{\nu}_e$. Hence, the process can occur.

The above reaction is verified experimentally.

40.4 Stability of Proton

The laws of baryon and lepton number conservations are responsible for the stability of the proton. Let us consider a possible mode of a free proton decay :

$$p \rightarrow e^{+} + \pi^{0}$$

For this $B = 1 \rightarrow 0 + 0$

and $L = 0 \rightarrow (-1) + 0$

The total values of both B and L are *not* the same before and after the decay, and so the decay is forbidden by both B and L conservations. Proton decay is really forbidden because there are no baryons of mass smaller than that of proton.

40.5 Resonance Particles

Almost all of the known particles are unstable. They decay, spontaneously falling apart into several other particles. The only absolutely stable particles are the electron, the proton, the photon, and the neutrinos.

The decay of a kaon and that of a lambda particle are some examples of decays. Each of these particles lives for about 10^{-10} s from the instant of production to the instant of decay. By the standards of high-energy physics, a time interval of 10^{-10} s is a rather long time interval. If a particle lives that long, physicists can do experiments on it and with it. In the terminology of high-energy physics, any particle with a lifetime of 10^{-10} or even 10^{-14} s is regarded as stable.

However, during the past decade or so there has been a profusion of new particles. These new particles are characterised by a life time of about 10^{-23} s showing that they are extremely unstable. In order to avoid answering the question whether they are really fundamental particles or not, physicists have started calling them '*resonances*' or resonance particles. These are all strongly interacting. The word resonance has been chosen to indicate them because they are recognised by resonance peaks in the normal energy spectrum of an event.

Production. The short-lived particles cannot be detected directly, but their existence can be deduced from circumstantial evidence. Consider the case of the short-lived lambda particle Λ (1520), found in collisions between kaons and protons. If a beam of negative kaons impacts on the protons in a bubble chamber, a variety of reactions will occur. Two of these are :

$$K^- + p \to \Sigma^0 + \pi^0 \qquad ...(1)$$

$$K^- + p \to \Lambda + \pi^+ + \pi^- \qquad ...(2)$$

The reaction cross-section for each of the reactions is some function of energy. Fig. 40.4 shows how the cross-sections vary as function of the centre-of-mass energy. There is a local maximum in the cross-sections at a centre-of-mass energy of 1520 MeV. At this energy *all* the cross-sections have a peak. So, all the maxima are due to the same cause. In each case, the colliding particles merge to create a new particle :

$$K^- + p \to \Lambda\,(1520) \qquad ...(3)$$

This new particle subsequently decays into one or another of the pairs of particles that emerge from the collision :

$$K^- + p \to \Lambda\,(1520) \to \Sigma^0 + \pi^0 \qquad ...(4)$$

$$\to \Lambda + \pi^+ + \pi^- \qquad ...(5)$$

According to Eq. (3), the production of a lambda will occur only if the energy of the kaon plus the mass-energy of the proton add up to the mass-energy of the lambda.

The lambda particle in the reactions (4) and (5) plays a role analogous to that of the compound nucleus in nuclear reactions. Short-lived particles such as the lambda are often called *resonances*. Thus one speaks of the Λ (1520) as a $K^- p$ resonance.

Decay Time Calculation. The plots of Fig. 40.4 show that the cross-section is not only large when the centre-of-mass energy is exactly 1520 MeV, but is also quite large at energies somewhat below and somewhat above 1520 MeV. This indicates that there is an uncertainty in the energy for the production of a lambda. The magnitude of this uncertainty is about ± 16 MeV. The uncertainty in the production energy implies an uncertainty in the mass-energy of the lambda. The mass is uncertain by ± 16 MeV/c^2.

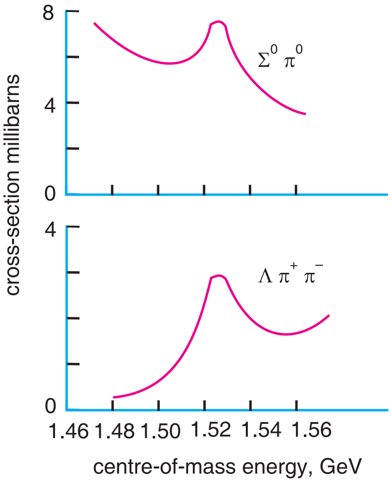


Fig. 40.4

$$M_{\Lambda\,(1520)} = 1520\ \text{MeV}/c^2 \pm 16\ \text{MeV}/c^2.$$

We can estimate the lifetime from the uncertainty in energy :

$$\Delta t \simeq \frac{\hbar}{\Delta E} \simeq \frac{1 \cdot 05 \times 10^{-34}\ \text{Js}}{16\,\text{MeV}} \simeq 4 \times 10^{-23}\ \text{s}$$

Detection of Neutrino.

Such extremely short life-times can be estimated only from the uncertainty principle, since an unstable particle does not travel far enough to leave a measurable track in a detector, say a bubble chamber, before decaying.

40.6 Detection of Neutrino

Reines and Cowan made use of the abundant supply of antineutrinos emitted from a nuclear reactor. Energetic antineutrinos can be detected by the reaction

$$\bar{\nu} + p \rightarrow n + \beta^+$$

The experimental arrangement is shown in Fig. 40.5.

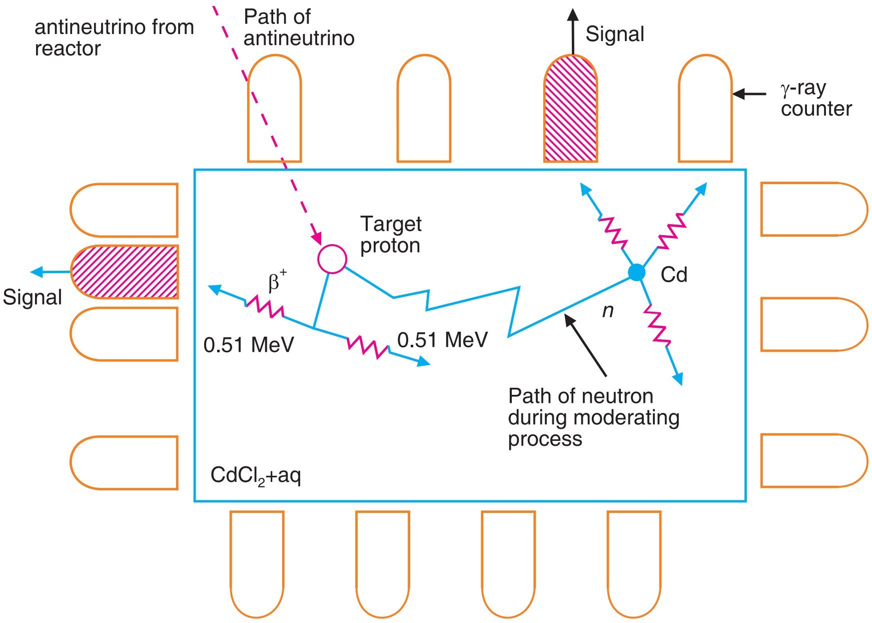

antineutrino from reactor
Path of antineutrino
Signal
γ-ray counter
Target proton
Signal
β+
Cd
n
0.51 MeV
0.51 MeV
Path of neutron during moderating process
CdCl2+aq

Fig. 40.5

An intense beam of antineutrinos, proceeding from a nuclear reactor is allowed to traverse a large plastic tank filled with an aqueous solution of cadmium chloride, $CdCl_2$. The tank is surrounded by many photomultiplier tubes. Suppose an antineutrino is absorbed by a proton and this system decays by β^+ emission to become a neutron. Thus

$$_1H^1 + \bar{\nu} \rightarrow {}_0n^1 + {}_1e^0.$$

The positron combines with an electron to create two γ photons, which are detected by the photomultiplier tubes. The neutron, moderated, *i.e.*, slowed down by collisions with protons, is finally captured by a cadmium nucleus when three γ photons are emitted. Thus,

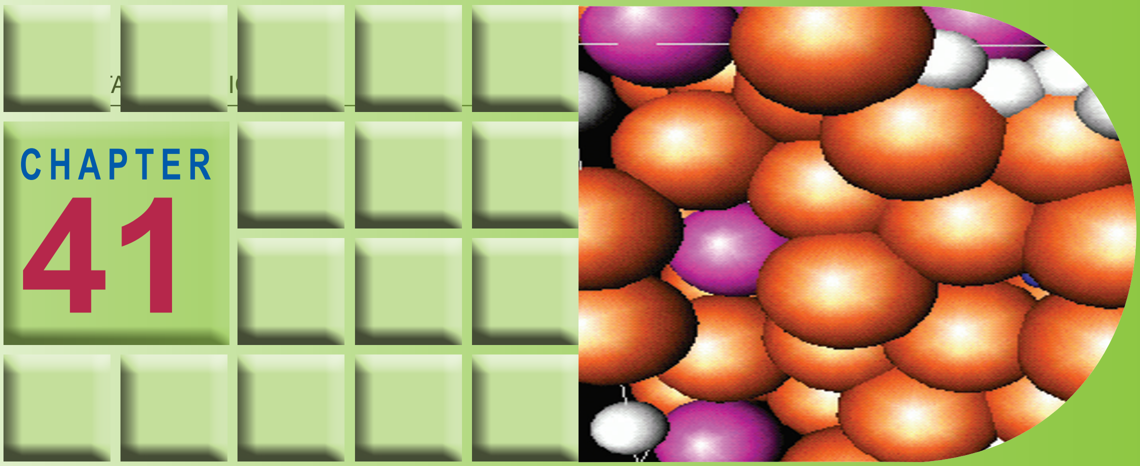
$$_{48}Cd^{113} + {}_0n^1 \rightarrow {}_{48}Cd^{114} + 3\gamma.$$

The almost simultaneous detection of these γ photons by the tubes established, beyond all reasonable doubt, the existence of antineutrinos.

PART – VI

SOLID STATE PHYSICS

41. Solid State Physics–I
42. Solid State Physics–II
43. Liquid Crystals
44. Superconductivity–I
45. Superconductivity–II
46. Reciprocal Lattice, Crystal Binding, Energy Bands
47. Magnetic Properties of Materials
48. Magnetism
49. Quantum Theory of Paramagnetism
50. Lattice Vibrations and Phonons
51. Electrical Conductivity and Ohm's Law
52. Free Electron Model of Metals
53. Free Electron Fermi Gas
54. Dielectric Properties of Solids



CHAPTER 41

SOLID STATE PHYSICS-I

AT A GLANCE

41.1	Bonding in Crystals	**41.2**	Ionic Bond
41.3	Covalent Bond	**41.4**	Metallic Bond
41.5	Molecular Bond	**41.6**	Hydrogen Bond
41.7	Calculation of Repulsive Exponent n	**41.8**	Born-Haber Cycle
41.9	Some Simple Crystal Structures	**41.10**	Specific Heat Capacity of Solids
41.11	Einstein's Theory of Specific Heat Capacity		
41.12	Debye's Theory of Specific Heat Capacity of a Solid		
41.13	Superconductivity		
41.14	The Meissner Effect		
41.15	The BCS Theory		

41.1 Bonding in Crystals

Most solids are crystalline, with the atoms, ions or molecules of which they are composed falling into regular, repeated three-dimensional patterns. The presence of *long-range order* is thus the defining property of a crystal. Crystals may be classified in terms of the dominant type of chemical binding force keeping the atoms together. All these bonds involve electrostatic forces, with the chief differences among them lying in the ways in which the outer electrons of the structural elements are distributed. The distinct types of bonds that provide the cohesive forces in crystals can be classified as follows: (*i*) the ionic bond (*ii*) the covalent bond (*iii*) the metallic bond (*iv*) the van der Waals bond and (*v*) the hydrogen bond. We briefly discuss the different types of bonds in crystals.

41.2 Ionic bond

Ionic bonds are formed when atoms that have low ionization energies, and hence lose electrons readily, interact with other atoms that tend to acquire excess electrons. The former atoms give up electrons to the latter. Thus the atoms become positive and negative ions respectively. These ions come together in an equilibrium configuration in which the attractive forces between positive and negative ions predominate over the repulsive forces between similar ions.

Consider the case of sodium chloride which is a typical example of an ionic crystal. Here, a single valence electron is transferred from the sodium atom to the chlorine atom. The Na^+ and Cl^- ions so formed are arranged in a face-centred cubic structure [Fig. 41.1 (*a*)]. A different arrangement is found in cesium chloride crystals. The body-centred cubic structure of a CsCl crystal is shown in Fig. 41.1 (*b*).

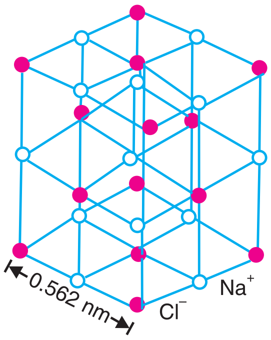


Fig. 41.1 (a)

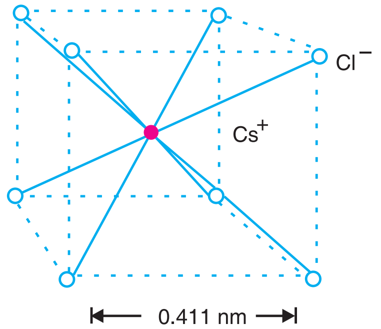
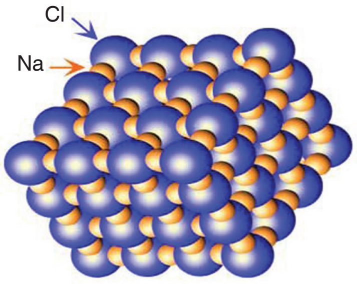


Fig. 41.1 (b)

Expression for the cohesive energy of an ionic crystal

The *cohesive energy* of an ionic crystal is the energy that would be liberated by the formation of the crystal from individual neutral atoms. The principal contribution to the cohesive energy of an ionic crystal is the electrostatic potential energy $U_{coulomb}$ of the ions. Let us consider an Na^+ ion in NaCl. Its nearest neighbours are six Cl^- ions, each one the distance r away.



Sodium Chloride Crystal Structure.

$$\left.\begin{array}{l}\text{The potential energy of the Na}^+\text{ ion}\\ \text{due to the 6 Cl}^-\text{ ions}\end{array}\right\} U_1 = -\frac{6e^2}{4\pi\varepsilon_0 r}$$

The next nearest neighbours are 12 Na^+ ions, each one the distance $\sqrt{2}\,r$ away.

$$\left.\begin{array}{l}\text{The potential energy of the Na}^+\text{ ion}\\ \text{due to the 12 Na}^+\text{ ions}\end{array}\right\} U_2 = +\frac{12e^2}{4\pi\varepsilon_0 \sqrt{2}\,r}$$

Then there are 8 Cl^- ions at $\sqrt{3}\,r$ distance, 6 Na^+ ions at $2r$ distance and so on. When the summation is continued over all the + and – ions in a crystal of infinite size, the result is

$$U_{\text{coulomb}} = -\frac{6e^2}{4\pi\varepsilon_0 r} + \frac{12e^2}{4\pi\varepsilon_0 \sqrt{2}\,r} - \frac{8e^2}{4\pi\varepsilon_0 \sqrt{3}\,r} + \frac{6e^2}{4\pi\varepsilon_0\, 2r} + \ldots$$

$$= -\frac{e^2}{4\pi\varepsilon_0 r}\left(\frac{6}{\sqrt{1}} - \frac{12}{\sqrt{2}} + \frac{8}{\sqrt{3}} - \frac{6}{\sqrt{4}} + \ldots\right)$$

$$= -1.748 \frac{e^2}{4\pi\varepsilon_0 r}$$

or in general, $$U_{\text{coulomb}} = -\alpha \frac{e^2}{4\pi\varepsilon_0 r} \qquad ...(1)$$

This result holds for the potential energy of a Cl^- ion also. α is called the *Madelung constant* of the crystal. It has the same value for all crystals of the same structure. For simple crystal structures α lies between 1.6 and 1.8.

Two ions cannot continuously approach each other under coulomb attraction on account of the exclusion principle. When they are at a certain small distance apart, they begin to repel each other with a force which increases rapidly with decreasing internuclear distance r. The potential energy contribution of the short-range repulsive forces can be expressed approximately in the form

$$U_{\text{repulsive}} = \frac{B}{r^n} \qquad ...(2)$$

where B is a constant and n is a number ($n \approx 9$).

Therefore, the total potential energy U of each ion due to its interactions with all the other ions is

$$U = U_{\text{Coulomb}} + U_{\text{repulsive}} = -\frac{\alpha e^2}{4\pi\varepsilon_0 r} + \frac{B}{r^n} \qquad ...(3)$$

We must now evaluate the constant B. At the equilibrium separation r_0 of the ions, U is a minimum. So $(dU/dr) = 0$ when $r = r_0$.

$$\left(\frac{dU}{dr}\right)_{r = r_0} = \frac{\alpha e^2}{4\pi\varepsilon_0 r_0^2} - \frac{nB}{r_0^{n+1}} = 0$$

or $$B = \frac{\alpha e^2}{4\pi\varepsilon_0 n} r_0^{n-1} \qquad ...(4)$$

The total potential energy at the equilibrium separation is

$$U = -\frac{\alpha e^2}{4\pi\varepsilon_0 r_0} + \frac{\alpha e^2 r_0^{n-1}}{4\pi\varepsilon_0 n r_0^n}$$

$\therefore$ $$U = -\frac{\alpha e^2}{4\pi\varepsilon_0 r_0}(1 - \frac{1}{n}) \qquad ...(5)$$

This is the magnitude of the energy needed to separate an ionic crystal into individual ions (not into atoms).

In an NaCl crystal, the equilibrium distance r_0 between ions is 2.81×10^{-10} m, $\alpha = 1.748$ and $n \approx 9$, $1/4\pi\varepsilon_0 = 9 \times 10^9$ Nm^2 C^{-2}.

The potential energy of an ion of either sign is

$$U = -\frac{\alpha e^2}{4\pi\varepsilon_0 r_0}(1 - \frac{1}{n})$$

$$= -\frac{(9 \times 10^9 \text{ Nm}^2 \text{ C}^{-2})(1.748)(1.60 \times 10^{-19} C)^2}{2.81 \times 10^{-10} \text{ m}}(1 - \frac{1}{9})$$

$$= -1.27 \times 10^{-18} \text{ J} = -7.96 \text{ eV}.$$

Because we may not count each ion more than once, only half this potential energy, or – 3.98 eV, represents the contribution *per ion* to the cohesive energy of the crystal.

Some energy is needed to transfer an electron from a Na atom to a Cl atom to form a $Na^+ - Cl^-$ ion pair. This electron transfer energy is the difference between the + 5.14 eV ionization energy of Na and the – 3.61 eV electron affinity of Cl, or + 1.53 eV. Each atom thus contributes + 0.77 eV to the cohesive energy.

$\therefore$ the total cohesive energy per atom in the NaCl crystal is

$$E_{cohesive} = (-3.98 + 0.77) \text{ eV/atom} = -3.21 \text{ eV/atom.}$$

Properties : (*i*) Most ionic solids are hard, brittle and have high melting points.

(*ii*) They are soluble in polar liquids like water.

(*iii*) Their electrical conductivity is much smaller than that of metals at room temperature. But in contrast to metals, the conductivity of ionic crystals increases with increasing temperature. At elevated temperatures, the ions themselves become mobile and ionic conductivity results.

(*iv*) Ionic solids crystallize in close-packed structures, of which the NaCl and CsCl structures are the commonest.

41.3 Covalent Bond

In the covalent bond, atoms are held together by the sharing of electrons. Each atom participating in a covalent bond contributes an electron to the bond. These electrons are shared by both atoms rather than being the virtually exclusive property of one of them as in an ionic bond. Diamond is an example of a crystal whose atoms are linked by covalent bonds. Fig. 41.2 shows the structure of a diamond crystal. The tetrahedral arrangement is a consequence of the ability of each carbon atom to form covalent bonds with four other atoms.

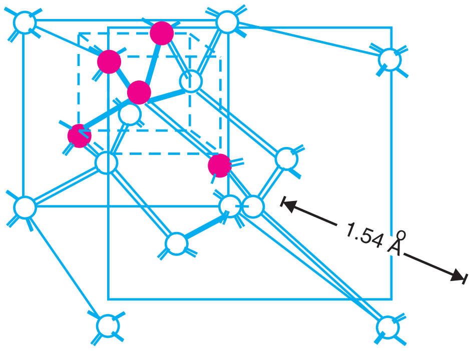


Fig. 41.2

The binding of molecular hydrogen (H2) is a simple example of covalent bond. In the H2 molecule, two electrons are shared by the two atoms (Fig. 41.3). As these electrons circulate, they spend more time between the atoms (in fact between the protons) than elsewhere and this produces an attractive force. Thus covalent crystals are composed of neutral atoms having slightly overlapping electron clouds. Diamond, germanium, silicon, and silicon carbide (SiC) are examples of covalent crystals. In SiC each atom is surrounded by four atoms of the other kind in the same tetrahedral structure as that of diamond.

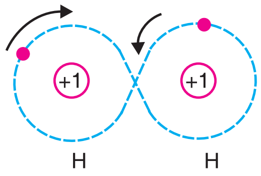
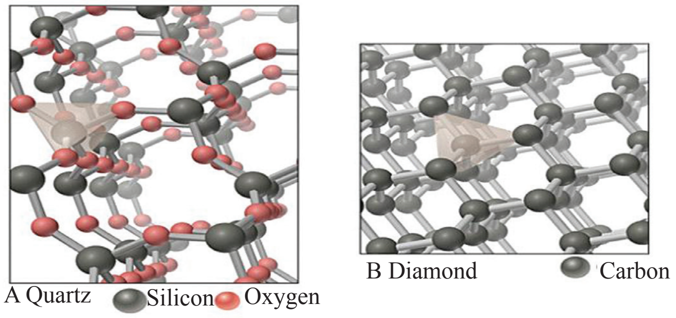


Fig. 41.3

Characteristics of Covalent Crystals:

(1) Covalent bond is a strong bond. Cohesive energies of 6 to 12 eV/atom are typical of covalent crystals, which is more than the usual cohesive energies in ionic crystals. All covalent crystals are hard, have high melting points, and are insoluble in all ordinary liquids.



A Quartz

B Diamond

Covalent Bond Strongest Bond Formed by Sharing Electrons.

(2) Covalent bonds are *strongly directional.*

(3) The other important characteristic property of the covalent bond is its *saturability*. Saturability means that each atom can form covalent bonds only with a limited number of its neighbours.

(4) The conductivity of covalent crystals varies over a wide range. Some crystals are insulators (diamond) and some are semiconductors (*Ge*). The conductivity increases with the increase of temperature.

(5) The optical properties of the covalent crystals are characterized by high refractive index and high dielectric constant. Covalent crystals are transparent to long-wavelength radiation but opaque to shorter wavelengths.

41.4 Metallic Bond

In metallic crystals, the metallic bond arises when all of the atoms share all of the valence electrons. The valence electrons of the atoms comprising a metal are common to the entire aggregate, so that a kind of "*gas*" of free electrons pervades it. The crystal is held together by the electrostatic attraction between the negative electron gas and the positive metal ions. The best example for a metallic crystal is sodium.

The cohesion (*i.e.*, the ability to remain a solid) of the metallic crystal results from a combination of forces :

(*i*) the attraction of the electron cloud for the ion cores,

(*ii*) the mutual repulsion of the electrons, and

(*iii*) the mutual repulsion of the ion cores.

Fig. 41.4 shows a sketch of a reasonable form for the potential (free) energy contribution, as a function of nuclear spacing, from each of these charge interactions.

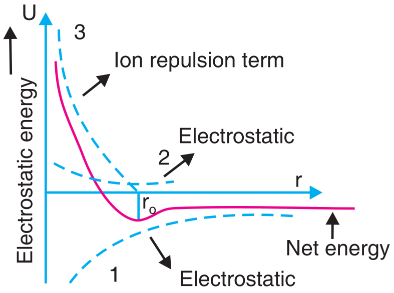


Fig. 41.4

The following are the characteristics of metal crystals:

(1) The presence of free electrons accounts for the high electrical and thermal conductivities of metals. The high electrical conductivity of metals is in turn responsible for their high optical reflection and absorption coefficients.

(2) The other characteristic properties of metals are their ductility and metallic lustre. Since the metallic bonds are not localised between adjacent atoms, the atoms of a metal can be rearranged in position without rupturing the crystal. This explains ductility of metals. When light shines on a metal, the free electrons oscillate under the electromagnetic field of the incident light and become sources of light. This gives the metal its surface lustre.

(3) Metallic bonds are weaker than ionic and covalent bonds.

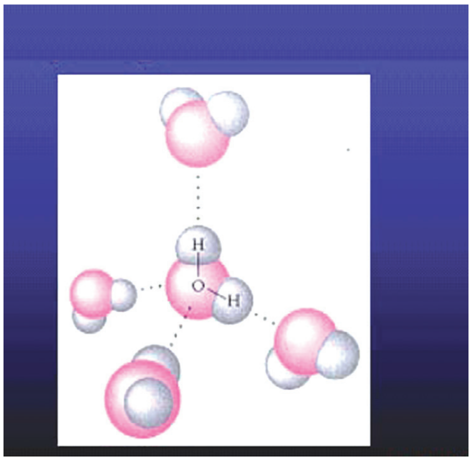
41.5 Molecular Bond

Neutral atoms with closed electron shells are bound together weakly by the *van der Waals forces*. The van der Waals' attraction was first explained for electrically neutral gas molecules by Debye. He assumed that neighbouring molecules induced dipoles in each other because of their own changing electric fields. This interaction produces an attractive force that is inversely proportional to the seventh power of the separation. The molecules are located at the crystal lattice points and the bonds between them are developed by van der Waals' forces. Solid argon and solid methane (CH_4) are examples of molecular crystals.

Properties: Van der Waals' forces are much weaker than those found in ionic and covalent bonds. As a result, molecular crystals generally have low melting and boiling points and little mechanical strength. Molecular crystals have low cohesive energies. Molecular crystals are good insulators due to non-availability of free electrons.

41.6 Hydrogen Bond

Hydrogen bond is formed under certain conditions when a single hydrogen atom appears to be bonded to two distinct electronegative atoms. This configuration is represented as $X - H \ldots Y$ where X is called 'donor' while Y is called 'acceptor'. The weaker of the two bonds (shown by dotted line) is the hydrogen bond while the other (shown by a full line) is a strong covalent bond. H_2O (ice), NH_3 and HF are examples of hydrogen-bonded crystals. These bonds are stronger than van der Waals bonds but weaker than ionic or covalent bonds.



Hydrogen Bond in Water

41.7 Calculation of Repulsive Exponent n

Born determined the repulsive exponent (n) from measurements of the compressibility of the crystal as follows:

The compressibility C (*i.e.*, reciprocal of bulk modulus K) is defined as

$$\frac{1}{C} = K = -V\left(\frac{dP}{dV}\right) \qquad \ldots(1)$$

where V is volume, P is pressure.

From the first law of thermodynamics

$$dQ = dU + P\,dV$$

At absolute zero there is no absorption of heat energy by the ions.

$$\therefore \qquad dQ = 0.$$

$\therefore$ change in internal energy, $dU = -P\,dV$

$$\frac{dU}{dV} = -P$$

or

$$\frac{d^2U}{dV^2} = -\frac{dP}{dV} \qquad \ldots(2)$$

$$\therefore \qquad \frac{1}{C} = V\frac{d^2U}{dV^2} \qquad \ldots(3)$$

Also

$$\frac{dU}{dV} = \frac{dU}{dr}\cdot\frac{dr}{dV} \quad i.e., \frac{d}{dV} = \frac{dr}{dV}\cdot\frac{d}{dr} \qquad \ldots(4)$$

and

$$\frac{d^2U}{dV^2} = \frac{d}{dV}\left(\frac{dU}{dV}\right) = \frac{d}{dV}\left(\frac{dU}{dr}\cdot\frac{dr}{dV}\right)$$

$$= \frac{dU}{dr}\cdot\frac{d^2 r}{dV^2} + \frac{dr}{dV}\cdot\frac{d}{dV}\left(\frac{dU}{dr}\right)$$

$$= \frac{dU}{dr}\cdot\frac{d^2 r}{dV^2} + \left(\frac{dr}{dV}\right)\cdot\frac{dr}{dV}\frac{d}{dr}\left(\frac{dU}{dr}\right)$$

$$= \frac{dU}{dr}\cdot\frac{d^2 r}{dV^2} + \left(\frac{dr}{dV}\right)^2\cdot\frac{d^2 U}{dr^2} \qquad \text{...(5)}$$

For NaCl crystal the volume of unit cell is

$$(2r)^3 = 8r^3$$

The unit cell has 4 sodium ions and 4 chloride ions. The volume $8r^3$ corresponds to 8 ions. So each ion corresponds to the volume r^3. If the crystal contains N–Na and N– Cl atoms, then $V = 2Nr^3$.

$$\therefore \qquad \frac{dV}{dr} = 6Nr^2 \text{ or } \frac{dr}{dV} = \frac{1}{6Nr^2}$$

$$\frac{d^2 r}{dV^2} = \frac{d}{dV}\left(\frac{dr}{dV}\right) = \frac{dr}{dV}\cdot\frac{d}{dr}\left(\frac{dr}{dV}\right) = \frac{1}{6Nr^2}\times\frac{d}{dr}\left(\frac{1}{6Nr^2}\right) = -\frac{1}{18N^2r^5}$$

Substituting these values in Eq. (5),

$$\frac{d^2 U}{dV^2} = -\frac{1}{18N^2r^5}\frac{dU}{dr} + \frac{1}{36N^2r^4}\frac{d^2 U}{dr^2} \qquad \text{...(6)}$$

At $r = r_0$, U is minimum and $dU/dr = 0$.

$$\left(\frac{d^2 U}{dV^2}\right)_{r=r_0} = \frac{1}{36N^2 r_0^4}\left(\frac{d^2 U}{dr^2}\right)_{r=r_0} \qquad \text{...(7)}$$

Substituting the above value in Eq. (3),

$$\frac{1}{C} = \left(V\frac{d^2U}{dV^2}\right)_{r=r_0} = 2Nr_0^3 \times \frac{1}{36N^2r_0^4}\left(\frac{d^2U}{dr^2}\right)_{r=r_0}$$

$$= \frac{1}{18Nr_0}\cdot\left(\frac{d^2 U}{dr^2}\right)_{r=r_0} \qquad \text{...(8)}$$

But

$$U = NU_i = N\left[\frac{B}{r^n} - \frac{\alpha e^2}{4\pi\varepsilon_0 r}\right] \qquad \text{...(9)}$$

$$\therefore \qquad \frac{dU}{dr} = N\left[-\frac{Bn}{r^{n+1}} + \frac{\alpha}{4\pi\varepsilon_0}\times\frac{e^2}{r^2}\right]$$

When

$$r = r_0, \frac{dU}{dr} = 0$$

$$-\frac{Bn}{r_0^{\,n+1}} + \frac{\alpha}{4\pi\varepsilon_0}\frac{e^2}{r_0^{\,2}} = 0$$

$$\therefore \qquad B = \frac{\alpha e^2 r_0^{\,n-1}}{4\pi\varepsilon_0 n} \qquad \text{...(10)}$$

Also $$\frac{d^2 U}{dr^2} = N\left[\frac{n(n+1)B}{r_0^{n+2}} - \frac{2\alpha e^2}{4\pi\varepsilon_0 r_0^3}\right]$$

Substituting the value of B,

$$\left(\frac{d^2 U}{dr^2}\right)_{r=r_0} = \frac{N\alpha e^2 (n-1)}{4\pi\varepsilon_0 r_0^3} \qquad ...(11)$$

Substituting this value in Eq. (8),

$$\frac{1}{C} = \frac{1}{18 N r_0} \frac{N\alpha e^2 (n-1)}{4\pi\varepsilon_0 r_0^3}$$

$$\therefore \qquad n = 1 + \frac{(4\pi\varepsilon_0) 18 r_0^4}{C\alpha e^2} \qquad ...(12)$$

Thus the value of n may be determined.

41.8 Born-Haber Cycle

Lattice energy : *The lattice energy is defined as the energy released in the process when the constituent ions are placed in their respective positions in the crystal lattice or this is the amount of energy which is spent to separate the solid ionic crystal into its constituent ions.*

An experimental check on the calculated values of the lattice energies may be obtained from Born-Haber cycle. The Born-Haber cycle for NaCl is represented as follows [Fig. 41.5] :

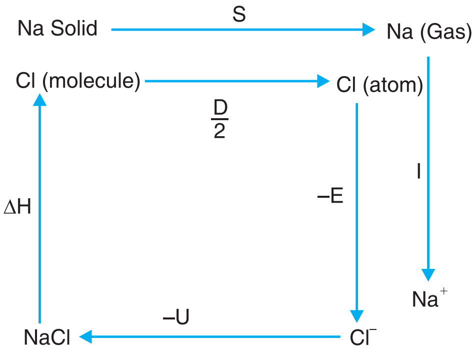


Fig. 41.5

Here S = Heat of sublimation, D = Dissociation energy

I = Ionisation energy, E = Electron affinity

ΔH = Heat of dissociation. U = Lattice energy.

The cycle can be explained as follows :

(*i*) Initially we start with solid sodium and chlorine molecule. The solid sodium is subjected to the sublimation energy (S) and it is vaporised. To the chlorine molecule, *dissociation energy* (D) is supplied and as a result it is dissociated into its constituent atoms. Here $\frac{D}{2}$ is the dissociation energy per chlorine atom.

(*ii*) Next ionisation energy I is supplied. The outer electron of Na gaseous atom is removed. This electron is added to Cl atom. As the Cl atom has the electron affinity E, an energy E is given out.

(*iii*) The two ions (Na^+ and Cl^-) are arranged in the lattice and hence the *lattice energy* U is released.

(*iv*) We again reach the starting point by supplying an energy (ΔH) known as *heat of dissociation.*

Now $$S + \frac{D}{2} + I - E - U + \Delta H = 0$$

$$\therefore \qquad U = S + \frac{D}{2} + I - E + \Delta H$$

All quantities excepting U are experimentally determinable. Hence the lattice energy U can be calculated.

41.9 Some Simple Crystal Structures

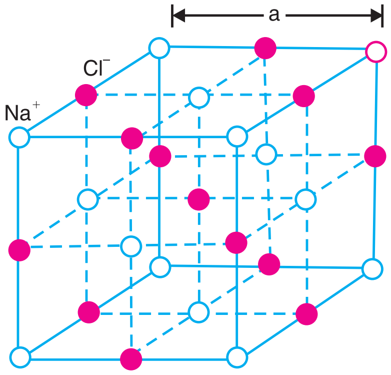


Fig. 41.6

NaCl Structure. The NaCl crystal has a fcc structure with Na^+ and Cl^- ions as the basis. This is an ionic crystal.

Fig. 41.6 represents a unit cell of NaCl lattice. The Na^+ ions are situated at the corners as well as at the centres of the faces of the cube. The Cl^- ions are relatively displaced half the edge of the unit cell along each axis. Thus NaCl crystal can be thought of as composed of superposed fcc Na^+ and Cl^- sub-lattices.

Each cell has 8 corners and 8 cells meet at each corner. Thus an ion at a corner of the cell is shared by 8 cells. So only 1/8 ion belongs to any one cell. Similarly, an ion at the centre of a face of the cell is shared by 2 cells. So only 1/2 ion belongs to any one cell. A cell has 8 corners and 6 faces. So it has $\left(\frac{1}{8}\times 8\right)+\left(\frac{1}{2}\times 6\right)=4$ ions of one kind, and similarly 4 ions of the other kind. Thus *there are* 4 $Na^+ - Cl^-$ *ion pairs per unit cell.*

There are four molecules of NaCl in each unit cube, with ions in the positions :

$$\text{Na}: 000; \frac{1}{2}\frac{1}{2}0;\ \frac{1}{2}0\frac{1}{2};\ 0\frac{1}{2}\frac{1}{2}$$

$$\text{Cl}: \frac{1}{2}\frac{1}{2}\frac{1}{2}; 0\,0\,\frac{1}{2};\ 0\,\frac{1}{2}\,0;\ \frac{1}{2}\,0\,0$$

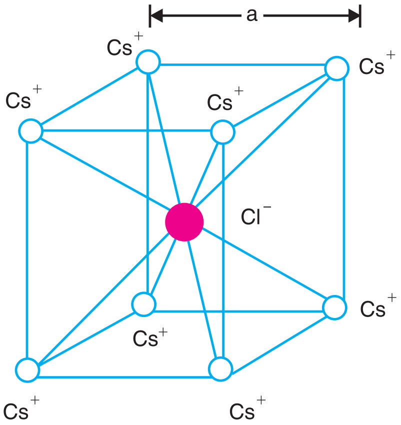
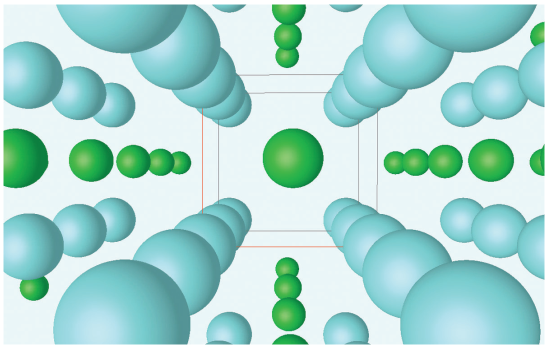


Fig. 41.7

Each Na^+ ion has 6 Cl^- ions as nearest neighbours. Similarly each Cl^- ion has 6 Na^+ ions as nearest neighbours. Hence the coordination number of NaCl for opposite kind of ions is 6 and the nearest neighbour distance is $a/2$. For same kind of ions coordination number is 12 and nearest neighbour distance is $a/\sqrt{2}$.

The Cesium Chloride Structure (CsCl). The structure of CsCl is shown in Fig. 41.7. The CsCl crystal has simple cubic structure with Cs^+ and Cl^- ions as the basis. The Cs^+ ions are situated at the corners of the cubic cell and Cl^- ion at the body centre or *vice versa.* Thus CsCl crystal may be considered to be a combination of two simple cubic sub-lattices, one of Cs^+ ion and the other of Cl^- ion. There is one molecule per primitive cell, with ions at the corners 0 0 0 and body-centred positions $\frac{1}{2}\frac{1}{2}\frac{1}{2}$ of the simple cubic space lattice. The coordination number for same kind of ions is 6



Cesium Chloride Structure

and the nearest neighbour distance is a. The coordination number for opposite kind of ions is 8 and the nearest neighbour distance is $\frac{a\sqrt{3}}{2}$.

Diamond Structure. The diamond structure is a fcc structure with a basis of two carbon atoms, one located at (0, 0, 0) and other at (a/4, a/4, a/4) associated with each lattice point. Fig. 41.8 shows the positions of atoms in the cubic cell of the diamond structure projected on a cubic face. The fractions denote height above the base in units of a cube edge. The points at 0 and $\frac{1}{2}$ are on the fcc lattice while those at 1/4 and 3/4 are on a similar lattice displaced along the body diagonal by one-fourth of its length. Thus the diamond lattice is composed of two inter-penetrating fcc sub-lattices, one of which is shifted relative to the other by one-fourth of a body diagonal.

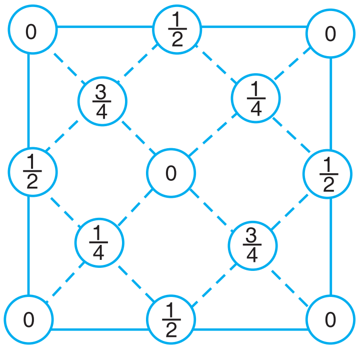


Fig. 41.8

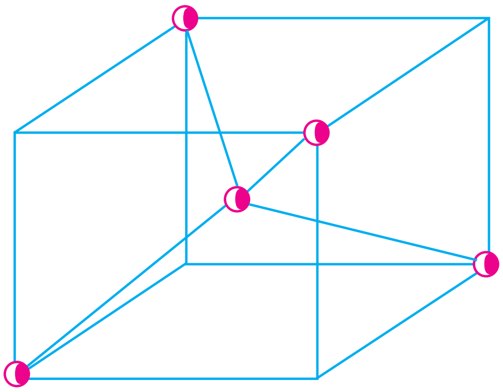


Fig. 41.9

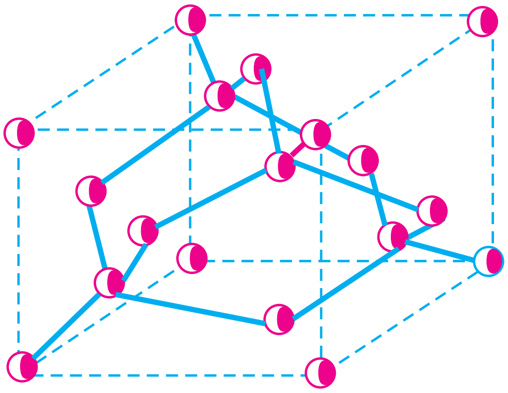
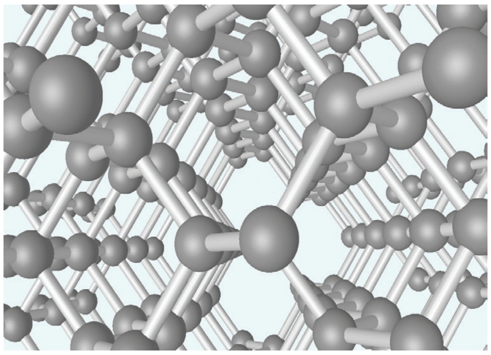


Fig. 41.10

In a diamond crystal the carbon atoms are linked by directional covalent bonds. Each carbon atom forms covalent bonds with four other carbon atoms that occupy four corners of a cube in a tetrahedral structure (Fig. 41.9). The length of each bond is 0.154 nm and the angle between the bonds is 109.5°. The entire diamond lattice is constructed of such tetrahedral units (Fig. 41.10).



Diamond Structure

In the diamond lattice, each atom has four nearest neighbours with which it forms covalent bonds. Thus the coordination number of diamond crystal is 4. The number of atoms per unit cell is 8. The diamond structure is loosely packed since each atom has only 4 nearest neighbours. Carbon, silicon, germanium and grey tin crystallize in diamond structure.

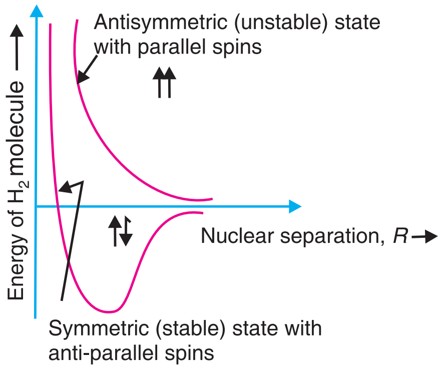


Fig. 41.11

The Hydrogen Molecule. The binding of molecular hydrogen is a simple example of a covalent bond. The bond length in H_2 is 0.074 nm. Fig. 41.11 shows the variation of the energy of the system H + H with their distances apart when the electron spins are parallel and antiparallel. The difference between the two curves is due to the Pauli exclusion principle, which leads to a dominating repulsion when the spins are parallel. The strongest binding occurs when the spins of the two electrons are anti parallel. The binding depends on the relative spin orientation not because there are strong magnetic dipole forces between the spins, but because the Pauli principle modifies the distribution of charge according to the spin orientation. A spin-dependent coulomb energy of the system is found. This is called the *exchange interaction.*

41.10 Specific Heat Capacity of Solids

The variation of atomic heat capacity with temperature for diamond, aluminium and silver is shown in Fig. 41.12.

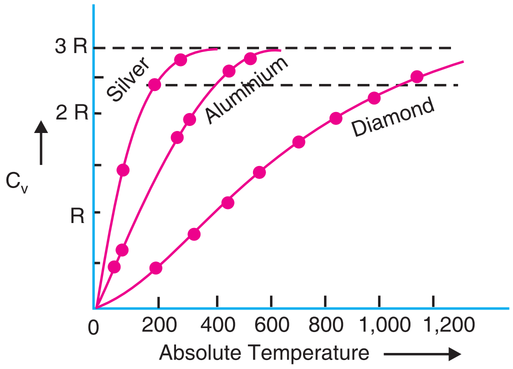
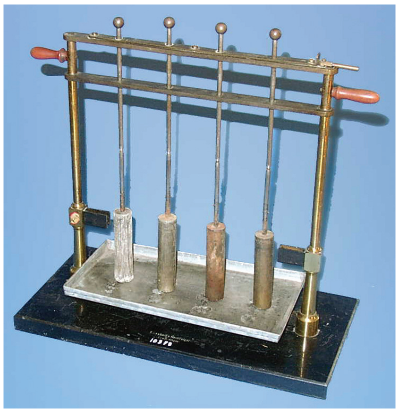


Fig. 41.12



Demonstration of Specific Heat Capacities of Solids

The variation of specific heat capacity with temperature could not be explained on the basis of the classical theories of heat. The first attempt to explain the variation of specific heat capacity with temperature was made by Einstein in 1907 on the basis of quantum theory.

41.11 Einstein's Theory of Specific Heat Capacity

According to quantum theory, heat is radiated from a body in the form of discrete particles, called *photons.* The energy of a photon is equal to $h\nu$ where h is the Planck's constant and ν is the frequency of radiation. Einstein assumed :

(1) Every solid is composed of atoms. The atoms are at rest at $0K$. The thermal energy is zero at $0K$. When the solid is heated, the atoms vibrate simple harmonically about their mean equilibrium positions. All the atoms vibrate with the same frequency.

(2) Each atom has three degrees of freedom.

(3) The mean energy per degree of freedom is $\dfrac{h\nu}{e^{h\nu/kT}-1}$ as calculated by Planck by the application of quantum theory.

The energy of each atom = $\dfrac{3h\nu}{e^{h\nu/kT}-1}$

$\therefore$ The internal energy of a solid consisting of N atoms

$$U = \frac{3N h\nu}{e^{h\nu/kT} - 1}$$

∴ Atomic heat capacity at constant volume

$$C_v = \frac{dU}{dT} = 3Nk\left(\frac{h\nu}{kT}\right)^2 \frac{e^{h\nu/kT}}{[e^{h\nu/kT} - 1]^2}$$

$$C_v = 3R\left\{\left(\frac{\theta}{T}\right)^2 \frac{e^{\theta/T}}{(e^{\theta/T} - 1)^2}\right\} \text{ where } \theta = \frac{h\nu}{k} \qquad \text{...(1)}$$

This is Einstein's equation for the atomic heat of a solid at constant volume.

Merits. (1) In the above equation, as $T \to O$, $C_v \to O$. Also according to Equation (1), atomic heat capacity decreases with decrease of temperature.

(2) At high temperatures, the atomic heat capacity approaches $3R$, which is in agreement with experiments.

Demerit. It was observed that at low temperature, there is some disagreement between experiment and theory. It was observed that in the case of some elements like copper, aluminium, iron etc., the atomic heat capacity at low temperature decreases more rapidly than that predicted by Einstein's theory. The disagreement is due to the assumption that all atoms of a solid vibrate at the same frequency. Debye considered the atoms, not as isolated oscillators, all vibrating at the same frequency, but as a system of coupled oscillators, having a continuous spectrum of natural frequencies. On this assumption, Debye modified Einstein's theory.

41.12 Debye's Theory of Specific Heat Capacity of a Solid

Debye assumed that :

(1) Any solid is capable of vibrating elastically in many different modes.

(2) The frequency of vibration in one mode is different from that in another mode.

(3) The number of modes of vibration of solids are limited in number.

A solid can be considered as an elastic body in which stationary waves, of both the longitudinal and transverse types, are set up. The frequencies range from 0 to a definite upper limit ν_m. This limit is chosen so that the total number of possible independent vibrations is equal to the number of vibrations of the separate atoms in the solid. The number of independent modes of vibration per unit volume, with frequencies between ν and $\nu + d\nu = 4\pi\left(\frac{1}{C_l^3} + \frac{2}{C_t^3}\right)\nu^2\, d\nu$.

Here, C_l = velocity of longitudinal waves and C_t = velocity of transverse waves.

If there are N atoms in volume V,

$$\left.\begin{array}{l}\text{The total No. of independent}\\ \text{modes of vibration}\end{array}\right\} = \int_0^{\nu_m} 4\pi V\left(\frac{1}{C_l^3} + \frac{2}{C_t^3}\right)\nu^2\, d\nu.$$

Each atom has 3 degrees of freedom and hence the possible number of vibrations will be equal to $3N$. Therefore,

$$\int_0^{\nu_m} 4\pi V\left(\frac{1}{C_l^3} + \frac{2}{C_t^3}\right)\nu^2\, d\nu = 3N \qquad \text{...(1)}$$

or

$$4\pi V\left(\frac{1}{C_l^3} + \frac{2}{C_t^3}\right) = \frac{9N}{\nu_m^3} \qquad \text{...(2)}$$

According to quantum theory, the energy associated with each degree of freedom is $\dfrac{h\nu}{e^{h\nu/kT}-1}$.

Hence, the energy in the solid for frequencies between ν and $\nu + d\nu$

$$= 4\pi V\left(\frac{1}{C_l^3}+\frac{2}{C_t^3}\right)\nu^2\,d\nu \times \frac{h\nu}{e^{h\nu/kT}-1}$$

$$= \frac{9N}{\nu_m^3}\frac{h\nu^3\,d\nu}{(e^{h\nu/kT}-1)} \qquad \text{...(3)}$$

∴ The total thermal energy associated with the solid

$$U = \frac{9N}{\nu_m^3}\int_0^{\nu_m}\frac{h\nu^3\,d\nu}{(e^{h\nu/kT}-1)} \qquad \text{...(4)}$$

Let $\dfrac{h\nu}{kT} = x$; then $\nu = \dfrac{kT}{h}x$ and $d\nu = \dfrac{kT}{h}dx$.

Let us introduce a characteristic temperature θ defined by $\theta = h\nu_m / k$.

When $\nu = 0$, $x = 0$ and when $\nu = \nu_m$, $x = \dfrac{h\nu_m}{kT} = \dfrac{\theta}{T}$.

Substituting these values in equation (4),

$$U = 9Nk\frac{T^4}{\theta^3}\int_0^{\theta/T}\frac{x^3}{e^x-1}dx \qquad \text{...(5)}$$

$$C_v = \frac{dU}{dT}$$

$$= 9R\left[4\left(\frac{T}{\theta}\right)^3\int_0^{\theta/T}\frac{x^3\,dx}{e^x-1}-\frac{\theta}{T}\frac{1}{e^{\theta/T}-1}\right] \qquad \text{...(6)}$$

This is Debye's formula for the atomic heat capacity.

Test of Debye's Theory. (1) At *high temperatures*, $x = h\nu/kT$ is small and θ/T also is small.

Hence, $e^{\theta/T} = 1+\theta/T$. Thus, $\dfrac{\theta}{T}\dfrac{1}{e^{\theta/T}-1} = 1$.

Also $e^x - 1 = x$. Hence,

$$4\left(\frac{T}{\theta}\right)^3\int_0^{\theta/T}\frac{x^3\,dx}{e^x-1} = 4\left(\frac{T}{\theta}\right)^3\int_0^{\theta/T}x^2\,dx = \frac{4}{3}.$$

Hence, $C_v = 9R\left(\frac{4}{3}-1\right) = 3R$.

This is in agreement with Dulong and Petit law and experiment.

(2) *At low temperatures*, x and θ/T are large. The exponential $e^{\theta/T}$ becomes large, so that the last term in equation (6) vanishes. The upper limit for integration can be taken as ∞ instead of θ/T. The value of the definite integral is $\dfrac{\pi^4}{15}$. Hence, at low temperatures,

$$C_v = 9R\times 4\left(\frac{T}{\theta}\right)^3\times\frac{\pi^4}{15} = \frac{12\pi^4}{5}R\left(\frac{T}{\theta}\right)^3$$

This is known as Debye's T^3 law. Hence, at low temperatures, the atomic heat is directly proportional to the cube of the absolute temperature. This is confirmed by experiment. The curve C_v against T is shown in Fig. 41.13.

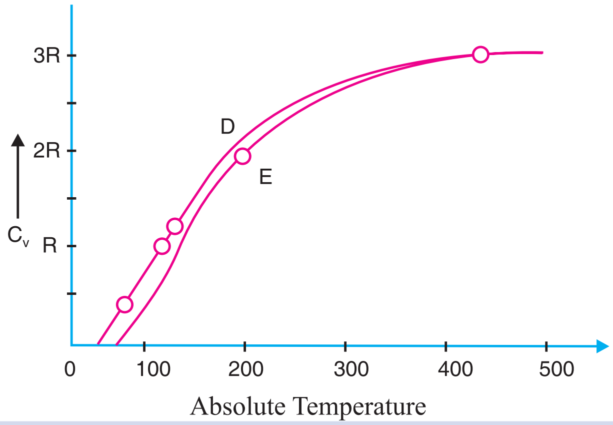


Fig. 41.13

41.13 Superconductivity

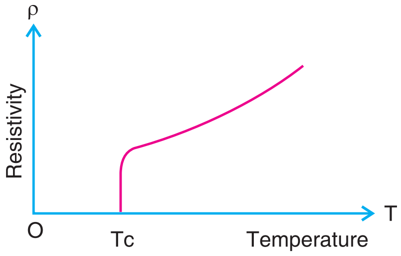
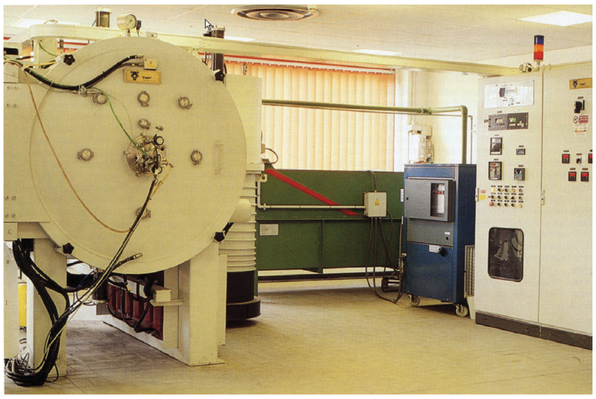


Fig. 41.14

Introduction. Superconductivity was first observed in 1911 by Dutch physicist H.K. Onnes in the course of his experiments on the electrical conductivities of metals at low temperatures. He observed that as purified mercury is cooled, its resistivity vanished abruptly at 4.2 K (Fig. 41.14). Above this temperature, the resistivity is small, but finite, while the resistivity below this point is so small that it is essentially zero. The temperature at which the transition takes place is called the critical temperature (T_c). The temperature (T_c), which marks the transition of a normal conductor to the superconducting state, is defined as the transition temperature. Above the critical temperature (T_c), the substance is in the familiar *normal state*, but below (T_c) it enters an entirely different *superconducting state*. The superconducting state is marked by a sudden fall of the electrical resistivity of the material to nearly zero, when it is cooled to a sufficiently low temperature.



Vacuum Furnace in Superconductivity Lab

We can gain some insight into the nature of superconductivity using the free electron model.

The resistivity of a metal may be written as $\rho = \dfrac{m}{ne^2\tau}$ where m = mass of the electron, e = charge of the electron, n = No. of electrons per unit volume and τ = collision time. The vibration of the ions in a crystal decreases with decreasing temperature. This has the effect of decreasing the probability of an electron-ion collision or of increasing the collision time τ. Thus ρ decreases as the temperature is lowered. According to the above equation, this implies that the resistivity of a metal should tend toward zero, as the temperature approaches zero. If τ becomes infinite at sufficiently low temperatures, then the resistivity vanishes entirely, which is what is observed in superconductivity. As the temperature is lowered below T_c, a fraction of the electrons become superconducting, in the sense that they have infinite collision times. These electrons undergo no scattering whatsoever, even though the substance may contain some impurities and defects. It is these electrons which are responsible for superconductivity.

Onnes found that the superconducting transition is *reversible* :

When he heated the superconducting sample, it recovered its normal resistivity at the temperature T_c.

Superconductivity occurs in metallic elements (silver, lead etc.) of the periodic system and also in alloys and semiconductors. The range of transition temperatures at present extends from 23.2 K for the alloy $Nb_3\ Ge$ to 0.01 K for some semiconductors.

Experimental facts. (1) When impurities are added to superconducting elements, the superconducting property is not lost but the transition-temperature is lowered.

(2) **Isotope effect.** It has been observed that the critical temperature of superconductors varies with isotopic mass. In mercury, T_c varies from 4.185 K to 4.146 K as the average atomic mass M varies from 199.5 to 203.4 atomic mass units. The experimental results within each series of isotopes may be fitted by a relation of the form $M^{\alpha}\ T_c$ = constant. Here α is a number usually 0.5. From the dependence of T_c on the isotopic mass M, we learn that lattice vibrations and hence electron-lattice interactions are deeply involved in superconductivity.

(3) The thermal properties such as the specific heat capacity and thermal conductivity of a substance change abruptly, when it passes over into the superconducting state.

(4) **Persistent currents.** If a superconductor has the form of a ring, current can be induced in it by electromagnetic induction. One usually measures the resistivity of a superconductor by observing this current as a function of time. If the sample is in the normal state, the current damps out quickly because of the resistance of the ring. But if the ring has zero resistance, the current, once set up, flows indefinitely without any decrease in value. In a typical experiment, a lead ring could carry an induced current of several hundred amperes for over a year without any change. Such currents are called 'persistent currents'. Physicists found that the upper limit for the resistivity of a superconducting lead ring was about 10^{-25} *ohm-m*. The fact that this is about $1/10^{+17}$ as large as the value at room temperature does indeed justify taking $\rho = 0$ for the superconducting state.

(5) **Destruction of superconductivity by Magnetic fields.** Superconductivity can be destroyed by the application of a magnetic field. If a strong enough magnetic field, called the *critical field*, is applied to a superconducting specimen, it becomes normal and recovers its normal resistivity even at $T < T_c$. The critical value of the applied magnetic field for the destruction of superconductivity is denoted by $H_c\ (T)$ and is a function of the temperature. For a given substance, value of H_c decreases as the temperature increases from $T = 0\ K$ to $T = T_c$. The variation can be represented by the formula

$$H_c(T) = H_c\,(0)[1 - (T/T_c)^2]$$

where H_c (0) is the critical field at 0 K. Thus the field has its maximum value, H_c (0) at $T = 0$ K. At the critical temperature (T_c) the critical field is zero : $H_c\ (T_c) = 0$. This result is expected, because at $T = T_c$ the specimen is already normal, and no field is necessary to accomplish the transition. The variation of the critical field with temperature for several superconducting elements is shown in Fig. 41.15. A specimen is superconducting below the curve and normal above the curve.

(6) The feature that we have been describing, that is, the destruction of superconductivity by relatively small magnetic fields, prevents the superconductor from being used in a solenoid to produce extremely large magnetic fields. However, there are superconductors that behave differently

from what we have described. Those that we have been describing are called type-I superconductors and they have two states: superconducting and normal. The other type of superconductor called type-II exists in three states: superconducting, mixed and normal. The mixed state is resistance-less but unlike the superconducting state flux from an applied magnetic field penetrates through it. That is, the Meissner effect does not occur in these type-II superconductors. In consequence, the mixed state persists up to high magnetic fields. Type-II superconductors can be used to wind solenoids capable of producing magnetic fields above 10 *T*.

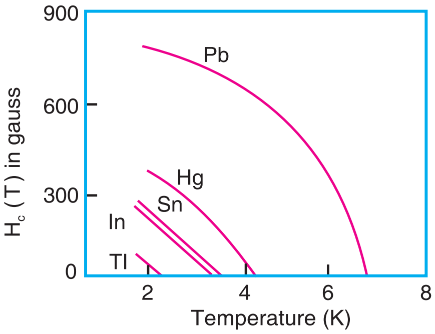


Fig. 41.15

41.14 The Meissner Effect

The Meissner Effect

Meissner and Ochsenfeld found that if a superconductor is cooled in a magnetic field to below the transition temperature, then at the transition, the lines of induction **B** are pushed out (Fig. 41.16).

The expulsion of magnetic flux from the interior of a piece of superconducting material as the material undergoes the transition to the superconducting phase is known as Meissner effect. In a series of experiments on superconducting cylinders, they demonstrated that, as the temperature is lowered to T_c, the flux is suddenly and completely expelled, as the specimen becomes superconducting, as shown in Fig. 41.16. The flux expulsion continues for all $T < T_c$. They established this by carefully measuring the magnetic field in the neighbourhood of the specimen. Further, they demonstrated that the effect is *reversible*: When the temperature is raised from below T_c, the flux suddenly penetrates the specimen after it reaches T_c and the substance is in the normal state.

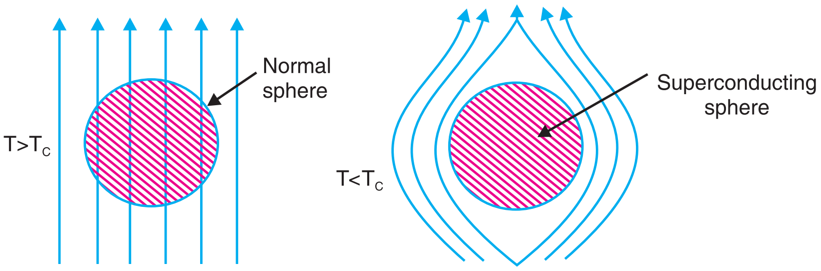


Fig. 41.16

A bulk superconductor behaves in an external magnetic field as if inside the specimen $B = \mu_0 (H + M) = \mu_0 (1 + \chi) H = 0$

(Here M = Magnetisation in the medium and χ its magnetic susceptibility).

Hence $\chi = -1$, *i.e.,* superconductor exhibits *perfect diamagnetism.*

41.15 The BCS Theory

The modern theory of superconductivity was promulgated by Bardeen, Cooper, and Schrieffer in 1957. It had as its major feature the pairing of electrons. Starting from first principles, and employing a completely quantum treatment, the BCS theory explains the various observable effects, such as zero resistance, the Meissner effect etc. In an ordinary metal, the electrical resistance is the result of the collisions of the conduction electrons with the vibrating ions in the crystal lattice. In the superconducting state, the forces of attraction between the conduction electrons exceed the forces of electrostatic repulsion.

Cooper pair. Normally two electrons repel each other, because of Coulomb interaction. Suppose that, for some reason, the two electrons attract each other. Cooper showed that the two electrons would then form a *bound* state. This is very important, because, in a bound state, electrons are paired to form a single system, and their motions are correlated. The pairing can be broken only if an amount of energy equal to the binding energy is applied to the system. This pair of electrons is called a *Cooper pair*. At temperature $T < T_c$, the lattice-electron interaction is stronger than electron-electron force of Coulomb. At this stage, the Cooper pairs of electrons will have a peculiar property of smoothly sailing over the lattice point without any energy exchange, *i.e.*, the Cooper pairs are not scattered by the lattice points. Hence no transfer of energy takes place from the electron pair to the lattice ions. If an electric field is established inside the substance, the electrons gain additional kinetic energy and give rise to a current. But they do not transfer this energy to the lattice, so that they do not get slowed down. As a consequence of this, the substance does not possess any electrical resistivity. So BCS theory explains the zero resistivity of a superconductor. The Cooper pair of electrons can maintain the coupled motion up to a certain distance among the lattice points in a superconductor called *Coherence length*. This is found to be of the order of 10^{-6} m.

We have been talking about the consequences of electron-electron attraction but how does this attraction come about in the first place? In super-conductive materials, it results from the electron-lattice interaction (Fig. 41.17). Suppose that the two electrons, 1 and 2, pass each other. Because electron 1 is negatively charged, it attracts positive ions toward itself (electron-lattice interaction). Thus electron 2 does not 'see' the bare electron 1. Electron 1 is screened by ions. The screening may greatly reduce the effective charge of this electron. In fact, the ions may over-respond and produce a net positive charge. If this happens, then electron 2 will be attracted toward electron 1. This leads to a net attractive interaction, as required for the formation of Cooper pair. In technical literature, one says that each electron is surrounded by a "phonon cloud", and that the two electrons establish an attractive interaction by exchanging phonons; for example, electron 1 emits phonons which are very quickly absorbed by electron 2, as in Fig. 41.17.

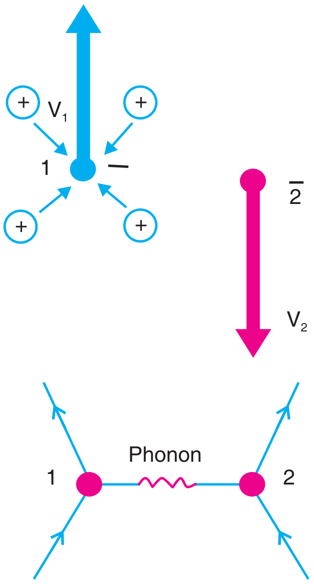
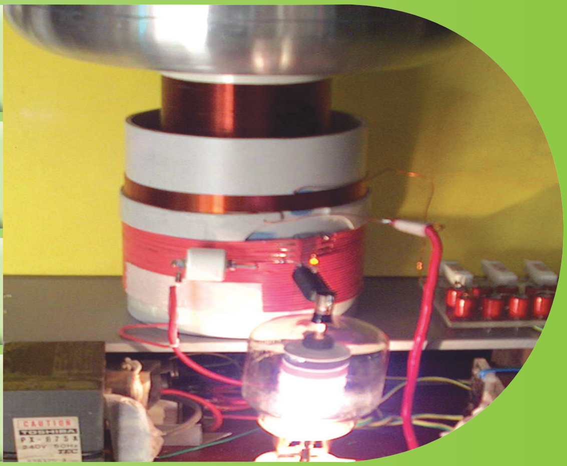


Fig. 41.17

CHAPTER 42



SOLID STATE PHYSICS–II

AT A GLANCE

42.1	Classification of Superconductors	**42.2**	Principle of Magnetic Levitation
42.3	X-ray Diffraction: Laue Equations	**42.4**	Magnetoresistance
42.5	Rotating Crystal Method	**42.6**	Powder Method (Debye-Scherrer Method)

42.1 Classification of Superconductors

The superconductors can be classified into two distinct groups according to their behaviour in external magnetic field:

1. Type-I Superconductors

The superconductors, in which the magnetic field is totally excluded from the interior of superconductors below a certain magnetising field H_C, *and at* H_C *the material loses superconductivity and the magnetic field penetrates fully are termed as type I or soft superconductors.*

Examples : Tin, Lead, Aluminium, Mercury, Zinc, Magnesium etc. The magnetization curve for type I superconducting materials is shown in Fig. 42.1.

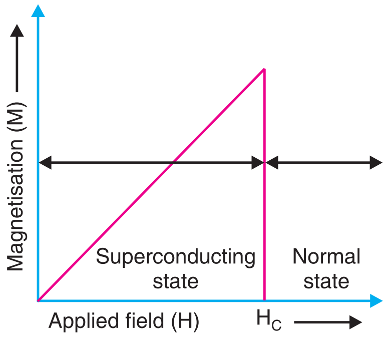


Fig. 42.1

The important characteristics of type I superconductors are:

(*i*) They exhibit complete Meissner effect.

(*ii*) The critical values of magnetic field H_C at which magnetisation drops are very low. The maximum known critical field for type I superconductor is of the order of 0.1 T. The low value of H_C makes these materials unsuitable for use in high field superconducting magnets.

The magnetisation curve shows that transition at H_C is reversible. This means that if the magnetic field is reduced below H_C, the material again acquires superconducting property and the field is expelled.

(*iii*) Below H_C the material is superconductor and above H_C it becomes a conductor.

2. Type II (or hard) Superconductors

The superconductors in which the material loses magnetisation gradually rather than suddenly are termed as type II or hard superconductors.

The magnetisation curve for these superconductors is shown in Fig. 42.2.

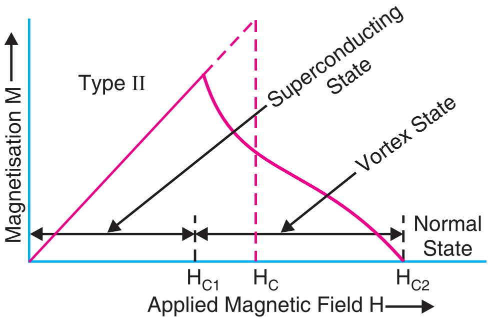


Fig. 42.2

The important characteristics are :

(*i*) They do not show complete Meissner effect.

(*ii*) These superconductors have two critical fields: H_{C1} (the lower critical field) and H_{C2} (the upper critical field). The specimen is diamagnetic below H_{C1}, *i.e.*, the magnetic field is completely excluded below H_{C1}. At H_{C1} the flux begins to penetrate the specimen, and the penetration of flux increases until the upper critical field H_{C2} is reached. At H_{C2} the magnetisation vanishes and specimen returns to normal conducting state. The value of H_{C2} is greater than H_C. The value of critical field for type II materials may be 100 times or more higher than the value of H_C for type I superconductors. Critical fields H_{C2} up-to 30 T have been observed.

The materials which display type II behaviour are essentially inhomogeneous and include Nb–Zr, Nb–Ti alloys and Va-Ga and Nb-Sn inter-metallic compounds. These are technically more useful than type I superconductors due to tolerating high magnetic fields.

42.2 Principle of Magnetic Levitation

The diamagnetic property of a superconductor is the basis of magnetic levitation. When a material is superconducting, the magnetic flux lines will be expelled from the material or the flux lines cannot penetrate the material. For example, if a small light and powerful magnet is kept over a superconducting material, the magnet will be **levitated** (lifted up) and it will float in air. This is known as *magnetic levitation.* It can be demonstrated using a high T_c superconductor like Y_1 $Ba_2Cu_3O_7$ compound taken in bulk and a light weight and powerful magnet (samarium-cobalt magnet). The Y-123 compound is cooled by pouring liquid nitrogen over it. The magnet kept over it just floats in air. In Japan, superconducting magnets have been used to levitate an experimental train above its rails and drive it at great speed.

42.3 X-ray Diffraction: Laue Equations

Let us examine X-ray scattering by two identical scattering centres in the crystal separated by a distance **r**. Let $\hat{n}_0$ be defined as unit vector in the direction of the incident beam and $\hat{n}_1$ a unit vector in an arbitrary scattered direction (Fig. 42.3). The incident radiation is assumed to be a parallel beam.

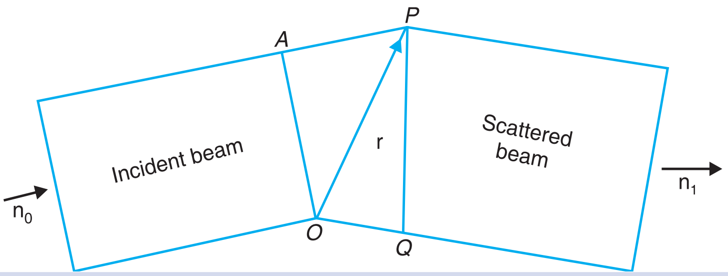


Fig. 42.3

The scattered beam is assumed to be detected at a very distant observation point.

The path difference between the waves scattered from the points *P* and *O* is

$$PA - OQ \;=\; \mathbf{r}.\hat{n}_0 - \mathbf{r}.\hat{n}_1 = \mathbf{r}.(\hat{n}_0 - \hat{n}_1) = \mathbf{r}.\mathbf{N} \qquad \text{...(1)}$$

Here, $\mathbf{N} = \hat{n}_0 - \hat{n}_1$. The vector **N** is the normal to Bragg reflecting plane whose magnitude is given by

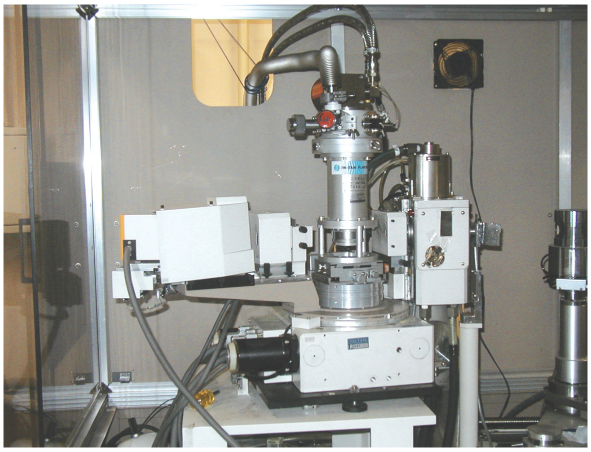
$$|\mathbf{N}| = 2\sin\theta \text{ (From Fig. 42.4)} \qquad \text{...(2)}$$

The phase difference ϕ_r between the radiation scattered at the two points is

$$\phi_r \;=\; \frac{2\pi}{\lambda}(\text{Path difference})$$

$$\;=\; \frac{2\pi}{\lambda}(\mathbf{r.N}) \qquad \text{...(3)}$$

The radiations scattered by atoms *P* and *O* will interfere constructively only if the phase difference becomes an integral multiple of 2π. Due to periodicity of the crystal, the other atoms placed in the same direction would also scatter the radiations in phase with those scattered from *P* and *O*. In a three-dimensional crystal, **r** may coincide with any of the three crystallographic axes **a, b** and **c**. Thus for the occurrence of a diffraction maximum, the following three conditions must be satisfied simultaneously:



X-ray Diffraction Equipment

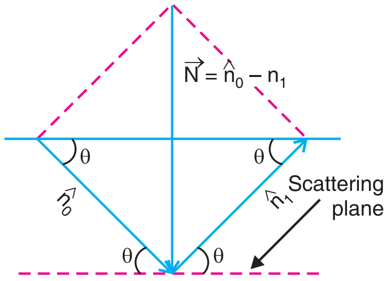


Fig. 42.4

$$\phi_a \;=\; \frac{2\pi}{\lambda}(\mathbf{a.N}) = 2\pi h' = 2\pi nh$$

$$\phi_b \;=\; \frac{2\pi}{\lambda}(\mathbf{b.N}) = 2\pi k' = 2\pi nk$$

$$\phi_c \;=\; \frac{2\pi}{\lambda}(\mathbf{c.N}) = 2\pi l' = 2\pi nl$$

Here, *h'*, *k'*, *l'* are any integers. *h'* = *nh*, *k'* = *nk*, *l'* = *nl*, where *h*, *k*, and *l* are three integers. Let α, β, γ be the angles between the scattering normal **N** and the crystallographic axes **a, b** and **c** respectively. Therefore, Eqs. (4) become

$$\begin{aligned}\mathbf{a.N} &= 2a \sin\theta \cos\alpha = h'\lambda = nh\,\lambda \\ \mathbf{b.N} &= 2b \sin\theta \cos\beta = k'\lambda = nk\,\lambda \\ \mathbf{c.N} &= 2c \sin\theta \cos\gamma = l'\lambda = nl\,\lambda \qquad ...(5)\end{aligned}$$

These equations are called *Laue equations*.

The three equations of (5) serve to determine a unique value for θ and **N**, thus defining a scattering direction. The direction cosines of the scattering normal **N** are seen from Eq. (5) to be proportional to $\frac{h}{a}, \frac{k}{b}$ and $\frac{l}{c}$. Neighbouring planes whose Miller indices are (*khl*) intersect the *a*, *b*, *c* axes at intervals $\frac{a}{h}, \frac{b}{k}, \frac{c}{l}$. The direction cosines of the normal to the (*h*, *k*, *l*) family of planes are therefore also proportional to $\frac{h}{a}, \frac{k}{b}$ and $\frac{l}{c}$. The scattering normal **N** is thus identical to the normal to the (*h*, *k*, *l*) planes and the (*h*, *k*, *l*) planes may be regarded as the reflecting planes of the Bragg picture.

If we put

$$h = \frac{a\cos\alpha}{d},\ k = \frac{b\cos\beta}{d},\ l = \frac{c\cos\gamma}{d}$$

each of these Laue equations reduces to Bragg's equation

$$2d \sin\theta = n\lambda.$$

Here *d* is the distance between adjacent planes of the (*hkl*) system.

42.4 Magnetoresistance

The change in electrical resistivity of a crystal when it is exposed to a magnetic field is called *magnetoresistance*. During the experiment on Hall effect, it is noticed that the resistance of the metal sample along the direction of current j_x generally increases (or the conductance decreases) after the magnetic field is switched on. The fractional increase in the zero-field resistivity ρ(0) is defined as the magnetoresistance Δρ/ρ(0). This effect of j_x is due to the fact that when the magnetic field is applied, the paths of the electrons become curved, and so the electrons do not go exactly in the direction of the superimposed electric field.

The magnetic resistance with one type of charge carrier, for small values of magnetic field *B* applied perpendicular to the electric field, is given by the expression

$$\frac{\Delta\rho}{\rho(0)} = \frac{\rho(B)-\rho(0)}{\rho(0)}\omega_c^2\tau^2 \qquad ...(1)$$

Here, ρ(0) is the zero field resistivity, ρ(*B*) is the resistivity in the presence of the magnetic field *B*, ω_c is the cyclotron frequency, eB/m^*, and τ is the collision time.

The relation (1) shows that $\Delta\rho \propto B^2$ in small magnetic fields.

At high fields, the magnetoresistance tends to saturate. Eq. (1) gives the transverse magnetoresistance.

The change in resistivity in the direction of applied magnetic field is called the *longitudinal magnetoresistance*. Magnetoresistance arises as all carriers are not moving with the same velocity in a combined electric and magnetic field, owing to their different effective masses. The origin of longitudinal magnetoresistance is attributed to the anisotropy in effective mass tensor or simply to the non-spherical shape of Fermi surfaces. The measurement of magnetoresistance in metal gives definite information about the topology of the Fermi surface.

EXPERIMENTAL METHODS IN X-RAY DIFFRACTION

42.5 Rotating Crystal Method

This method is widely used for the determination of the size of the unit cell of a crystal.

Apparatus. Fig. 42.5 (a) shows the experimental arrangement.

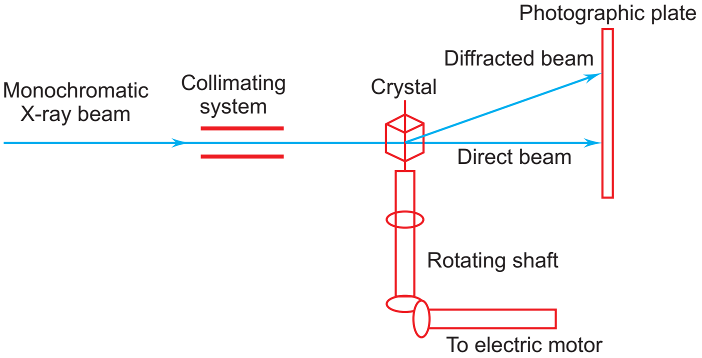

Photographic plate
Diffracted beam
Monochromatic X-ray beam
Collimating system
Crystal
Direct beam
Rotating shaft
To electric motor

Fig. 42.5 (a)

- *X*-rays are generated in the *X*-ray tube. The beam is made monochromatic by using a filter. The beam is then passed through the collimating system which gives a fine pencil of parallel *X*-rays.
- A small single crystal (of ~ 1 mm dimensions) is mounted on a shaft arranged perpendicular to the incident beam. The shaft is rotated by a small electric motor.
- The crystal is positioned at the centre of a cylindrical holder concentric with the rotating shaft [Fig. 42.5 (*b*)].
- A photographic film is attached at the inner circular surface of the cylinder.

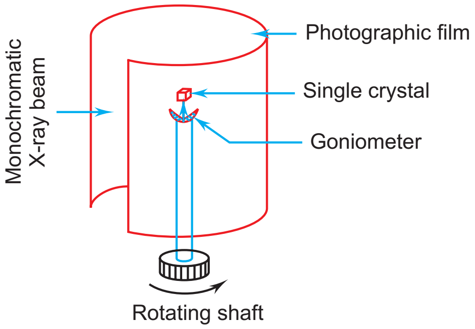

Photographic film
Single crystal
Goniometer
Monochromatic X-ray beam
Rotating shaft

Fig. 42.5 (b)

Experimental procedure

When the crystal is set into rotation about a fixed axis, the sets of atomic planes come successively into their reflecting positions. The diffracted beams from all planes parallel to the vertical rotation axis lie in the horizontal plane. The diffracted beams from planes with other orientations will lie in layers above and below the horizontal plane. The diffraction pattern is registered on a film of cylindrical camera, the axis of which coincides with the axis of rotation of the crystal.

Theory of rotation photograph

Assume that the crystal is mounted in such a way that one of the crystallographic axes coincides with the axis of rotation of the crystal, say *z*-axis. *B* and *E* are two neighbouring atoms at the lattice points along the z-axis (Fig. 42.6).

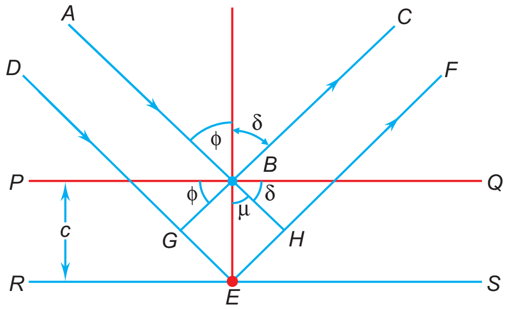


Fig. 42.6. Diffraction of X-rays by lattice elements along the rotation axis.

- The distance between B and E is c, the primitive translation.
- The angle of incidence and the angle of reflection are φ and δ respectively.
- G is the foot of perpendicular drawn from B on the incident ray DE.
- H is the foot of perpendicular drawn from B on the reflected ray EF.
- $\angle EBH = (90 - \delta) = \mu$.

The path difference between the two interfering rays is

$$GE + EH = BE \sin(90 - \varphi) + BE \sin(90 - \delta) = c(\cos\varphi + \cos\delta)$$

The condition of diffraction is given by

$$c(\cos\delta + \cos\varphi) = n\lambda \qquad \text{...(1)}$$

If the rotation axis is perpendicular to the beam, then $\varphi = 90°$ and Eq. (1) reduces to

$$c\cos\delta = c\cos(90 - \mu) = c\sin\mu = n\lambda \qquad \text{...(2)}$$

For various values of $n = 0, 1, 2, 3, ...$, we get the series of equations:

$$\cos\delta_0 = \sin\mu_0 = 0; \quad \cos\delta_1 = \sin\mu_1 = \frac{\lambda}{c};$$

$$\cos\delta_2 = \sin\mu_2 = \frac{2\lambda}{c}; \quad \text{... etc.}$$

$$\cos\delta_n = \sin\mu_n = \frac{n\lambda}{c} \qquad \text{...(3)}$$

Formation of layer lines on a cylindrical film

When λ is constant, such equations give the loci of all possible diffracted rays as the crystal turns. These loci are elements of a series of cones, of which the half apex angles are $\delta_0, \delta_1, \delta_2, \delta_3, ...$ etc. (Fig. 42.7). Any element of each of the cones makes an angle μ_0, μ_1, μ_2, etc., respectively with the horizontal plane.

The central horizontal plane contains all diffracted beams of order of zero. The Miller indices of all planes giving diffracted beams in horizontal plane must be represented by $(h, k, 0)$. Similarly, X-rays diffracted by planes of indices, (h, k, l) lie on the first order cone defined by

$$\mu_1 = \sin^{-1}\left(\frac{\lambda}{c}\right) \qquad \text{...(4)}$$

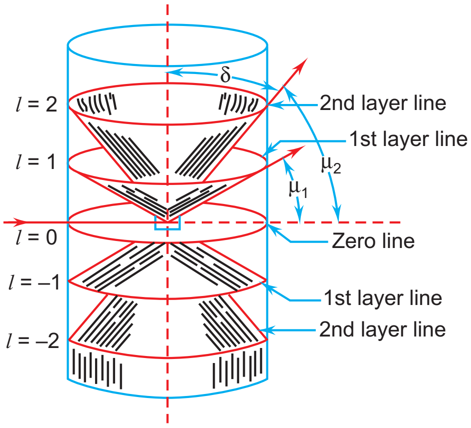


Fig. 42.7

In general, *X*-rays diffracted by planes of indices (*hkn*) lie on the nth order cone defined by

$$\mu_n = \sin^{-1}\left(\frac{n\lambda}{c}\right) \qquad ...(5)$$

Fig. 42.7 shows that the cones intersect the cylindrical film in a series of circles lying in planes perpendicular to the axis of rotation which becomes a series of parallel straight lines, known as **layer lines**, when the film is unrolled.

Fig. 42.8 shows layer *lines produced after flattening the photographic film.*

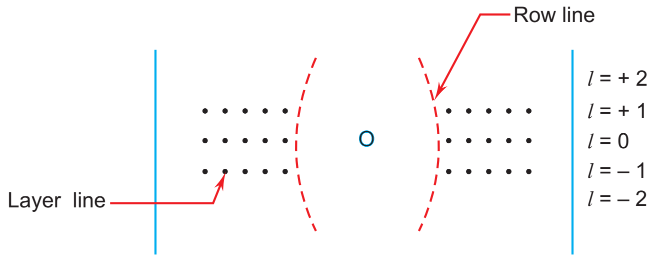


Fig. 42.8

- The cone for $n = 0$ is a plane perpendicular to the axis and containing the direction of incidence and the intersection of this direction with the film is known as *equatorial or zero layer line.* The succeeding lines are termed first, second, third, etc., layer lines.
- Diffraction images from planes with the same values of h and k but different values of l all lie on one of a series of curves called *row lines,* which are transverse to the layer lines.

The dimensions of the unit cell of the structure are determined in the following way: The spacing of layer lines give lattice translations. If the distance of the film from the crystal is known, the distance of the layer lines from the equatorial line gives the values of angle δ and as λ is known, the value of c can be calculated. Now the rotational photographs are taken separately about all three axes. In this way, lattice translations a, b, c are calculated which give the dimensions of the unit cell of the crystal.

42.6 Powder Method (Debye-Scherrer method)

Experimental arrangement: Fig. 42.9 shows the experimental arrangement.

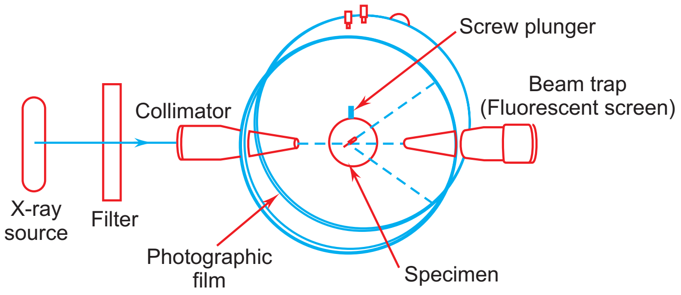


Fig. 42.9

- The finely powdered sample is filled in a thin capillary tube.
- The specimen is mounted on the holder at the centre of the cylindrical camera. The specimen is aligned using a screw plunger so that the axis of rotation of the specimen coincides with the axis of the camera. This ensures that the specimen is always within the *X*-ray beam profile during rotation.

- A collimated beam of monochromatic x-rays is produced by passing the x-rays through a filter and a collimator. The *x*-rays enter the camera through a fine hole in the collimator and fall on the polycrystalline rotating specimen. The direct beam being intense, enters a beam trap after passing through the specimen.

- The diffracted rays go out from individual crystallites which happen to be oriented with planes making a glancing angle θ with the beam satisfying Bragg's equation $2d \sin\theta = n\lambda$. The diffracted rays due to any one *hkl* plane leave the specimen along the surface of a cone concentric with the original beam with a semi-apex angle 2θ because of the spherical symmetry of the specimen (Fig. 42.10). In a similar way, other lattice planes generate their diffraction cones.

- These cones intercept the film strip, kept cylindrically along the inner wall of the camera, in a series of concentric rings (a pair of arcs).

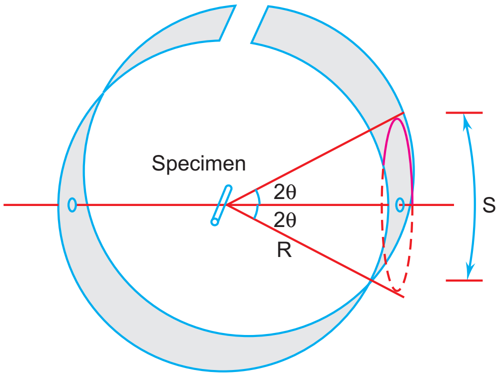


Fig. 42.10

- When the Bragg angle is 45°, the corresponding cone opens out into a circular disc which gives rise to a straight line at the midpoint of the film strip. When the Bragg angle is more than 45°, back reflections are produced. Thus, by using the cylindrical film, the Bragg angle from zero up to a maximum of 90° can be obtained.

Experimental procedure

- The sample is prepared, mounted on the sample holder and centred.

- The camera with the sample is taken to the dark room. Two holes are punched on the film. The film is loaded without disturbing the sample. The camera is closed, brought out and exposed to *X*-rays. During exposure, the sample is continuously rotated by connecting to a motor attachment. On completion of the exposure time (say 2 hrs for NaCl), the camera is taken to the dark room, the film is taken out, processed (developed and fixed) and dried for indexing.

Fig. 42.11 shows *flattened photographic film after developing and indexing of diffraction lines.*

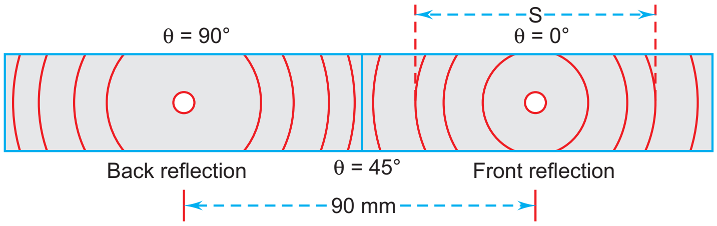


Fig. 42.11

Interpretation of the lines

- It is necessary to know the positions of low and high angle (θ = 0 and θ = 90°) diffraction lines on the film before any measurements are made. This can be done by observing the background intensity on the film. The background intensity is maximum near θ = 0. This arises due to the scattering of the *X*-rays by air molecules present inside the camera. Thus one expects more blackening of the film near the low angle side, *i.e.,* around the collimator.

- The film is placed over a comparator (film reader). The telescope is focussed on the lines. The linear distance (S) between different pairs of arcs is determined accurately. The radius (R) of the camera is known.

The angle θ corresponding to a particular pair of arcs is related to the distance S between the arcs as

$$4\theta \text{ (radians)} = S/R \qquad ...(1)$$

If θ is measured in degrees, Eq. (1) is modified as

$$4\theta \text{ (degrees)} = \frac{57.3\,S}{R} \qquad ...(2)$$

or

$$\theta = \frac{57.3\,S}{4R} \text{ degrees} \qquad ...(3)$$

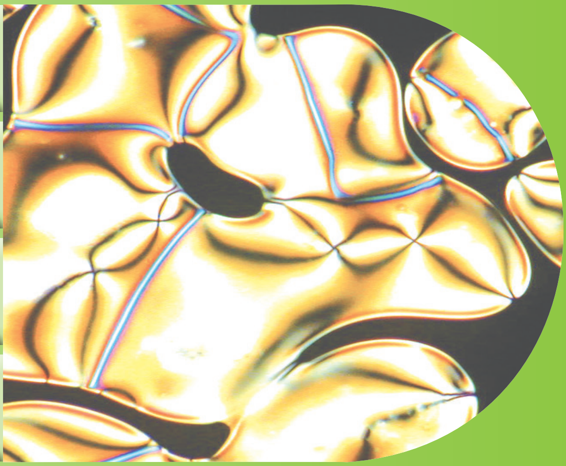
The various diffraction angles (θ) can be calculated using Eq. (3).

Knowing the Bragg angle θ for different lines, the interplanar spacings (d values) can be calculated using Bragg equation, *i.e.,* $2d \sin\theta = n\lambda$.

These d values are used to determine the space lattice of the crystal structure.

CHAPTER

43



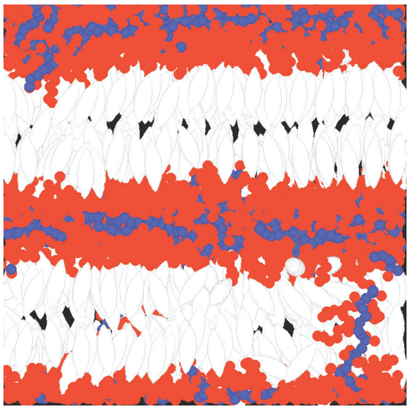
LIQUID CRYSTALS

AT A GLANCE

43.1	Liquid Crystals	43.2	Thermotropic Liquid Crystals
43.3	Lyotropic Liquid Crystals	43.4	Applications
43.5	Glass	43.6	Glass Transition Temperature
43.7	Metallic Glasses	43.8	Quasicrystals

43.1 Liquid Crystals

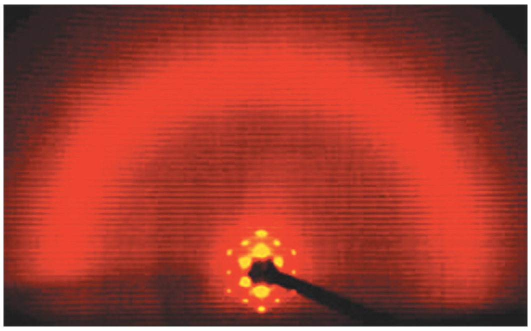
A liquid crystal is an organic compound which resembles to some extent both an ordered crystal and a disordered liquid. Liquid crystals have properties, some of crystals and some of liquids. Liquid crystals retain the anisotropy of properties such as dielectric, optical, magnetic and others, inherent in a crystal. They simultaneously acquire the properties specific to liquids — fluidity, ability to form droplets etc. The substances exhibiting such an intermediate, mesomorphic phase are known as *liquid crystals.*



Liquid Crystals

The temperature of transition from the solid phase to mesophase (liquid crystalline phase) is called the melting (or solidification) point T_s. *The temperature at which the liquid crystals convert to the isotropic transparent liquid is called the transparency temperature* T_t. *For different liquid crystals, the temperatures* T_s *and* T_t *and the temperature interval* $\Delta T = T_t - T_s$ *in which the liquid crystalline phase exists vary within wide limits.*

All liquid crystals are organic materials composed of rigid, moderately large rod-like molecules typically 1 to 3 nm in length and 0.3 to 1.0 nm across. At the end of molecules there are clusters of atoms exhibiting high polarizability. The ordered arrangement of molecules in the liquid crystalline state is provided by relatively weak van der Waals forces.



Thermotropic Liquid Crystals

Liquid crystal is not an isotropic medium. It is an anisotropic medium where the electric displacement vector **D** is not in the direction of electric vector **E.**

Liquid crystals are mainly classified into two types: (1) thermotropic and (2) Lyotropic.

43.2 Thermotropic Liquid Crystals

In thermotropic liquid crystals the molecular ordering changes with temperature. The molecules of liquid crystals have rod-like shape, with a central rigid portion and flexible ends. According to different types of molecular ordering, the thermotropic crystals are classified as nematics, cholesterics and smectics.

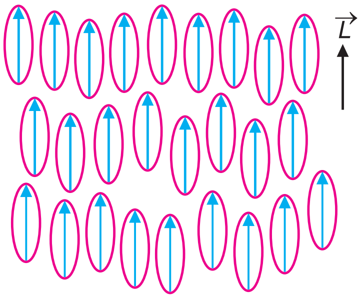
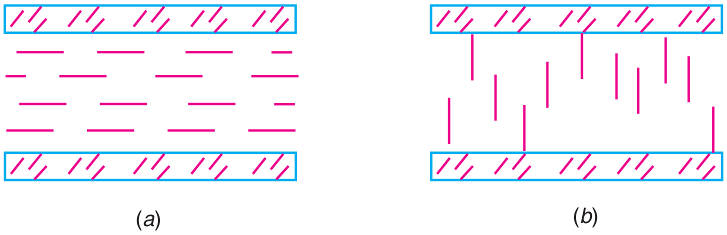


Fig. 43.1

Nematic Liquid Crystals

Nematic liquid crystals have rod-like molecules and the molecules have a natural tendency to orient their long axes along preferred direction called the director axis (Fig. 43.1). The entire nematic liquid crystal consists of small regions, each having its molecules aligned parallel to a unique axis, but this axis generally varies in direction in different regions of the crystal. This feature accounts for a high optical inhomogenity of the crystal, which appears opaque in the transmitted and reflected light.

An externally applied field orients the molecules of all regions in the same direction, and the liquid crystals become transparent. Placing a thin layer of the nematic liquid between two plate electrodes enables aligning the molecules both parallel, [Fig. 43.2 (*a*)] and normal, [Fig. 43.2 (*b*)], to the planes of the conducting plates. The first type of alignment is said to be *planar* and the second *hemotropic*. The latter structure, unlike the former,



(*a*) (*b*)

Fig. 43.2

has no effect on the polarization of light transmitted through the cell at a right angle to the layer of the liquid crystal. Special treatment of the surfaces of plate electrodes or requisite chemical aligning agents introduced in the liquid give the desired orientation of molecules. An electric field produced between the plates and various other factors can change this orientation and this alters the optical properties of the liquid crystal.

Pure nematic crystals are basically insulators. To make them conducting, they are to be doped with ionic impurities.

Cholesteric Liquid Crystals

The structure of a cholesteric liquid crystal is shown in Fig. 43.3. A cholesteric liquid crystal has its structure twisted about the helical axis lying perpendicular to the orientation of molecules. The vector **L** is called the director. In going from one plane to the other, the vector *L* turns by a certain angle and its end describes a helix with a pitch (structure period) of about 0.2 to 20 μm. The periodicity of the structure along the helical axis results in Bragg reflection of light at a wavelength equal to the pitch divided by the refractive index of the liquid crystal. A liquid crystal has high optical activity. So the cholesteric fluid causes the plane of polarization of light to turn through extremely large angles, of the order of 6,000 to 7,000° / mm. But in such an optically active crystal as quartz, the angle of turn of the light polarization plane is 15.5° / mm.

An external field applied to the cholesteric fluid can change the pitch of the helix, turn the helical axis, and convert the fluid to the nematic type. Therefore, the transmission of light by the system consisting of the layer of the cholesteric liquid placed between two polarizers will change with the applied potential difference. This is the principle that underlies the operation of various displays, liquid crystal digital indicators applied in some types of electronic watches, clocks, microcalculators, etc.

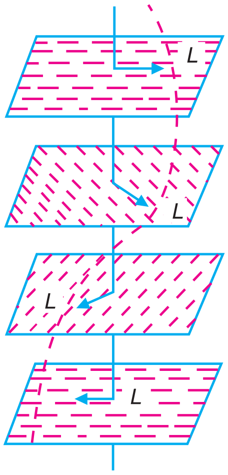
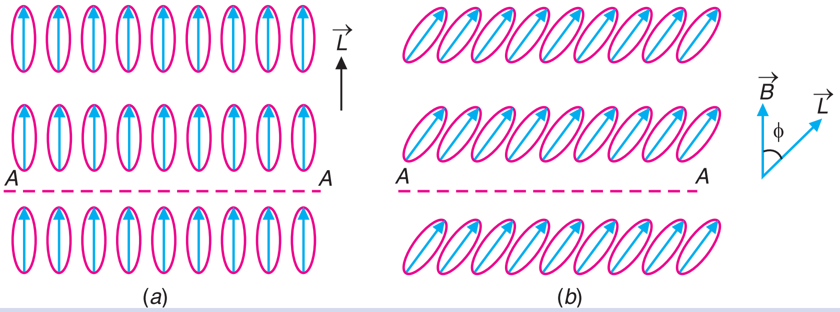


Fig. 43.3

Smectic Liquid Crystals

The most ordered smectic crystals have a layer structure. Molecules align themselves in parallel layers gliding one relative to the other, thereby promoting fluidity. The long axis of molecules line up either perpendicular to the planes of layers [Fig. 43.4 (*a*)] or form a certain angle ϕ with the normal to these planes, [Fig. 43.4 (*b*)]. The value of this angle ϕ depends on the temperature of liquid crystal and can be altered strongly by an externally applied field.

L is smectic axis of molecules and *AA* is shear plane.



(*a*) (*b*)

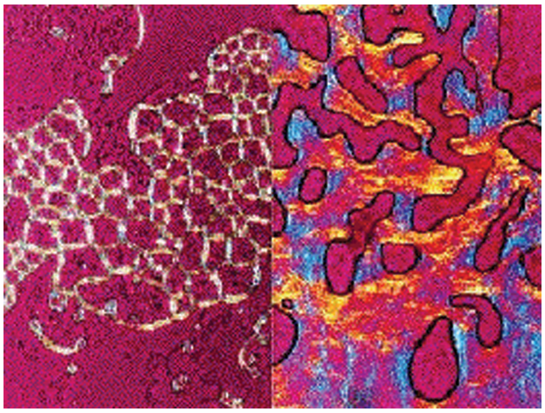
Fig. 43.4

43.3 Lyotropic Liquid Crystals

In this type, molecular ordering is changed by varying the concentration. Lyotropic liquid crystals are formed by amphiphilic molecules in suitable solvents with proper concentration.

Amphiphilic molecules have two parts :

(1) Water hating or hydrophobic part, and



Lyotropic Liquid Crystals

(2) Water loving or hydrophilic part.

Example. Stearic acid — $CH_3\,(CH_2)_{16}\,COOH$

Here $CH_3\,(CH_2)_{16}$ is the hydrophobic group and COOH is the hydrophilic group.

43.4 Applications

Liquid crystals have found wide application in the construction of low power visual displays, such as those found in watches and calculators. The action of the *LCD* is described in Section 56.5.

Liquid crystalline cells consume less power, are cheap and easy to manufacture in any size and shape and a variety of colours. But these cells also present some disadvantages, namely, they operate in a limited temperature range, have comparatively large response times, short life, etc.

43.5 Glass

Generally a *glass* is an amorphous, brittle and transparent solid. It is not metallic. It does not exhibit ferromagnetism. A glass has the elastic properties of an isotropic solid. A glass has the random structure of the liquid from which it is derived by cooling below the freezing point, without crystallization. Fig. 43.5 illustrates the difference between (a) a crystal and (b) a glass.

UNIVERSITY OF WISCONSIN
Eagles
1909
LA CROSSE

Glass does not exhibit Ferromagnetism

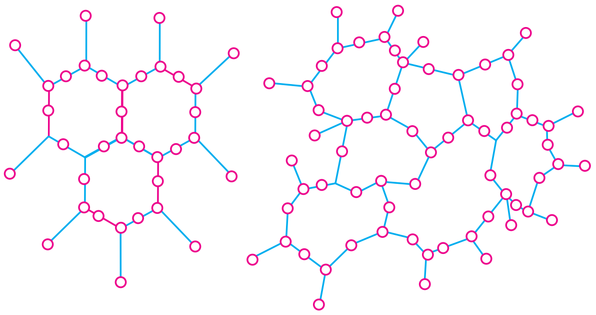

(*a*) (*b*)

Fig. 43.5

In certain cases, the solid state may correspond to a supercooled liquid in which the molecular arrangement of the liquid state is frozen. Because of rapid cooling and a high viscosity of the liquid, crystals may not have had time to grow and a glassy material results [Fig. 43.5(b)]. Upon annealing, such glassy substances may crystallize (devitrify).

The following Table illustrates the characteristics of glasses and compares these to those of metals.

Property	*Glass*	*Metal*
Structure	amorphous	crystalline
Bonding	covalent	metallic
Yield-stress	almost ideal	non-ideal
Workability	poor, brittle	good, ductile
Hardness	very high	low to high
Optical transmission	transparent	opaque
Thermal conductivity	poor	very good
Resistance	high	very low
Corrosion resistance	very good	poor to good
Magnetic properties	non-existent	various

43.6 Glass Transition Temperature

The glass transition temperature is the temperature at which the liquid like atomic configuration can be frozen into a solid.

A liquid on being cooled becomes a glass when the viscosity equals 10^{13} poise. This defines the *glass transition temperature* T_g. At temperatures above T_g we have a liquid. At temperatures below T_g we have a glass. The transition is not a thermodynamic phase transition, only a transition for practical purposes.

43.7 Metallic Glasses

Metallic glasses are metal alloys which have no long range atomic order. To achieve this structure, a suitable alloy must be quenched so rapidly that the random liquid like structure is retained in the solid.

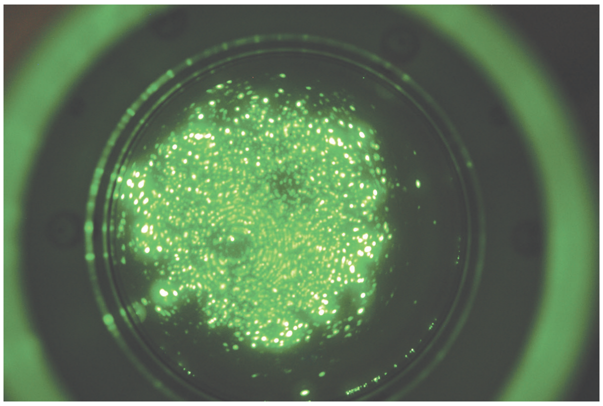
Properties

(*i*) Metallic glasses share the properties of both metals and glasses. They are amorphous alloys with an atomic configuration similar to that of the molten liquid, *i.e.*, there is no translational symmetry.

(*ii*) Metallic glasses are all strong, ductile, malleable, opaque and brittle.

(*iii*) A metallic glass has conductivity similar in magnitude to that of solid and liquid metals and optical behaviour typical of metals.

(*iv*) A metallic glass is homogeneous, non-crystalline, but disordered structure. It is densely packed as in a compact crystalline structure.



Mechanical Behavior of Metallic Glassy Films

Applications

(*i*) Even though metallic glasses are non-crystalline, they are ferromagnetic. They possess low magnetic losses, high permeability and saturation magnetisation with low coercivity. So they are used in tape-recorder heads, cores of high power transformers and magnetic shields.

EXAMPLE. $Fe_{75} P_{15} C_{10}$, $Fe_{24} Zr_{76}$, $Ni_{60} Nb_{40}$.

(*ii*) Metallic glasses possess high tensile strength and hardness. For example, the tensile strength of $Fe_{80}B_{20}$ is 3.6 GPa. For $Fe_{40} Mo_{40} B_{20}$, the hardness (HV) is 1950 kg/mm^2. So they are used as reinforcing elements in concrete, plastic or rubber.

(*iii*) The electrical resistance of metallic glasses is high and does not change appreciably with temperature. So they are used to make accurate standard resistances, computer memories and magneto resistance sensors.

In metallic glasses, there is a sharp variation in resistance at very low temperatures. So they are used as low temperature thermometers having resistivity of about 60 $\mu V/K$.

(*iv*) Many metallic glasses based on titanium, niobium, lanthanum and molybdenum are superconductors. So they are used in producing high magnetic fields for fusion reactors and for magnetic levitation of trains.

(*v*) Chromium and Phosphorus based metallic glasses have high corrosion resistance. So they are used in marine cables, chemical filters, inner surfaces of reactor vessels, orthopaedical implants and surgical clips.

43.8 Quasicrystals

We have learned that a necessary condition for having a crystalline phase is lattice periodicity and a crystal cannot have a five-fold rotational symmetry. But there are physical systems with many properties of the usual crystalline state but without three-dimensional lattice periodicity like $Cu_{0.2}\, A1_{0.65}\, Fe_{0.15}$. These systems do not have a strict crystal-like periodicity, but have only some subtle type of quasi/translational periodicity and a long-range order, but with five-fold symmetry axes. Such quasiperiodic crystals having many properties similar to usual familiar crystals are termed as *quasicrystals*. The full point symmetry group of the experimentally observed diffraction patterns of these five-fold symmetry crystals correspond to a platonic solid called an

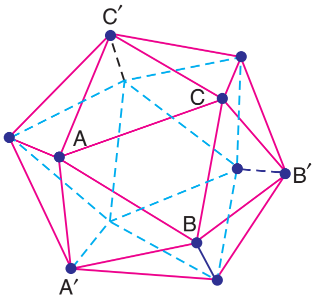


Fig. 43.6

icosahedron (Fig. 43.6). The icosahedron has six five-fold symmetry axes, 10 three-fold and 15 two-fold symmetry axes.

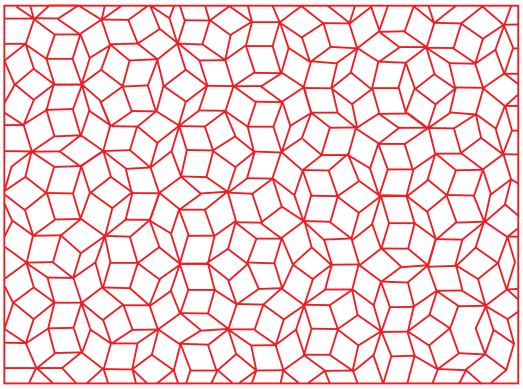


Fig. 43.7

Quasiperiodicity was first observed in grains of size 2 μm in the alloy of Al with 14 atomic per cent Mn. The smaller Mn atoms are each surrounded by 12 Al atoms arranged at the corners of an icosahedron. The structure is made up of parallel icosahedra attached at their edges. As a result of tiling in two dimensions, we get the structure known as *Penrose patterns* (Fig. 43.7). There is no long-range translationally periodic order, but there is a long-range orientational order: all of the space cannot be filled by repeating the basic unit, but a crystal can be constructed by nucleation at a centre cell, followed by an outward growth from there; thus overall, the crystal could not exhibit the five-fold symmetry of an icosahedron but the *X*-ray diffraction pattern could give a five-fold symmetry.

EXERCISE

1. What is a liquid crystal? Give an example.
2. What are the important characteristics of a liquid crystal ?

CHAPTER

44

SUPERCONDUCTIVITY–I

AT A GLANCE

44.1	The Josephson Effect	44.2	Theory of DC Josephson Effect
44.3	Theory of AC Josephson Effect	44.4	Flux Quantization
45.5	High-T_c Superconductivity	44.6	Applications

44.1 The Josephson Effect

In a Josephson junction two pieces of superconducting material are separated by a thin insulating oxide layer of thickness 2a (Fig. 44.1). The insulating layer is approximately 2 nm thick. The Cooper pairs can tunnel through the insulator, giving a current from one superconductor into the other. The tunnelling of a current through an insulating layer was discovered experimentally by I. Giaever. The behaviour of the current through an insulating layer sandwiched between two superconductors was first analyzed theoretically by B. D. Josephson. He predicted that a *'super current' consisting of correlated pairs of electrons can be made to flow across an insulating gap between two superconductors provided the gap is small enough.* Such a junction is called weak link. In practice, such a junction is manufactured by depositing a thin layer of an oxide, only 1 to 2 nm

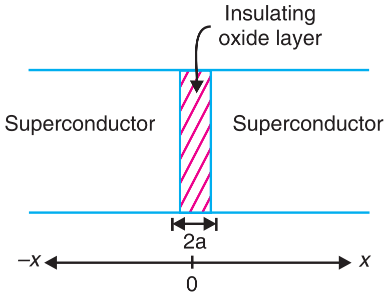


Fig. 44.1

thick, on the surface of one superconductor, and then placing the second superconductor on top of this.

The layer of insulator acts as a potential barrier for the Cooper pairs. A fraction of the current approaching this barrier from the left or the right manages to tunnel through the barrier and continue on the other side, without any net loss of energy. Thus, the junction permits the flow of a current even if the potential difference across it is zero. The phenomenon in which the junction permits the flow of current without any net loss of energy even if the potential difference across it is zero is called *DC Josephson effect.* It results from the familiar tunnelling phenomenon of quantum mechanics.

If we apply a DC voltage V across the junction, the result is an alternating current - the constant voltage generates an oscillating current. The frequency of this oscillating current is directly proportional to the voltage :

$$\nu = \frac{2eV}{h} \qquad \text{...(1)}$$

This is called the *AC Josephson effect.* It is caused by the coupling between the wave in one superconductor and the wave in the other. Crudely, it is analogous to the classical beat phenomenon seen in the superposition of two waves of different frequencies. The wavefunctions of a Cooper pair on the two sides of the junction differ in energy by 2eV; and hence differ in frequency by 2eV/h. Since the charge density and the current only depend on $|\psi|^2$, the individual wave frequencies are not observable. However, when the waves are coupled, the net amplitude is modulated by the beat frequency, which equals the frequency difference. The modulation shows up in the current.

According to Eq. (1), a voltage of 10^{-6} V generates an AC of a frequency of 483.6 MHz. Thus, the measurement of the frequencies of AC Josephson currents can be used as a very precise and convenient method for the measurement of voltage. Potentiometers based on this method attain a precision of about 1 part in 10^8. This effect has also been utilized for the precise determination of the value of $\frac{h}{e}$.

44.2 Theory of DC Josephson Effect

Let ψ_1 be the probability amplitude of electron pairs on one side of a junction and let ψ_2 be the amplitude on the other side. For simplicity, let us suppose that both superconductors are identical. Let us also suppose that they are both at zero potential. The time dependent Schrodinger equation

$$e\hbar\frac{\partial\psi}{\partial t} = H\psi$$

applied to the amplitudes ψ_1 and ψ_2 gives

$$i\hbar\frac{\partial\psi_1}{\partial t} = \hbar T\psi_2$$

and

$$i\hbar\frac{\partial\psi_2}{\partial t} = \hbar T\psi_1 \qquad \text{...(1)}$$

Here $\hbar T$ represents the effect of the electron-pair coupling or transfer interaction across the insulator. T has the dimensions of a frequency or rate. In fact, T is a measure of the leakage of ψ_1 into the region 2 and of ψ_2 into the region 1. If the insulator between the two superconductors is very thick, there is no pair tunnelling and T is zero.

Let

$$\psi_1 = n_1^{1/2}\, e^{i\theta_1} \quad \text{and} \quad \psi_2 = n_2^{1/2}\, e^{i\theta_2}$$

$$\frac{\partial\psi_1}{\partial t} = \frac{1}{2}n_1^{-1/2}\, e^{i\theta_1}\frac{\partial n_1}{\partial t} + i\psi_1\frac{\partial\theta_1}{\partial t} = -iT\psi_2 \qquad \text{...(2)}$$

and
$$\frac{\partial \psi_2}{\partial t} = \frac{1}{2} n_2^{-1/2} e^{i\theta_2} \frac{\partial n_2}{\partial t} + i\psi_2 \frac{\partial \theta_2}{\partial t} = -iT\psi_1 \qquad ...(3)$$

Multiplying Eq. (2) by $n_1^{1/2} e^{-i\theta_1}$ and using $\delta \equiv (\theta_2 - \theta_1)$, we get

$$\frac{1}{2}\frac{\partial n_1}{\partial t} + in_1 \frac{\partial \theta_1}{\partial t} = -iT(n_1 n_2)^{1/2} e^{i\delta} \qquad ...(4)$$

Multiplying Eq. (3) by $n_2^{1/2} e^{-i\theta_2}$, we get

$$\frac{1}{2}\frac{\partial n_2}{\partial t} + in_2 \frac{\partial \theta_2}{\partial t} = -iT(n_1 n_2)^{1/2} e^{-i\delta} \qquad ...(5)$$

Equating real and imaginary parts in Eqs. (4) and (5), we get

$$\left.\begin{aligned} \frac{\partial n_1}{\partial t} &= 2T(n_1 n_2)^{1/2} \sin\delta \\ \frac{\partial n_2}{\partial t} &= -2T(n_1 n_2)^{1/2} \sin\delta \end{aligned}\right\} \qquad ...(6)$$

$$\left.\begin{aligned} \frac{\partial \theta_1}{\partial t} &= -T\left(\frac{n_2}{n_1}\right)^{1/2} \cos\delta \\ \frac{\partial \theta_2}{\partial t} &= -T\left(\frac{n_1}{n_2}\right)^{1/2} \cos\delta \end{aligned}\right\} \qquad ...(7)$$

Taking $n_1 \cong n_2$ for identical superconductor 1 and 2, we have from Eq. (7),

$$\frac{\partial \theta_1}{\partial t} = \frac{\partial \theta_2}{\partial t}\ ;\ \frac{\partial}{\partial t}(\theta_2 - \theta_1) = 0 \qquad ...(8)$$

From Eq. (6), we have

$$\frac{\partial n_2}{\partial t} = -\frac{\partial n_1}{\partial t} \qquad ...(9)$$

The current flow through the junction *i.e.,* from 1 to 2 is proportional to $\frac{\partial n_2}{\partial t}$ or $\left(-\frac{\partial n_1}{\partial t}\right)$. Therefore, we can conclude that the *supercurrent* J of superconductor pairs across the junction depends on the phase difference δ, *i.e.,*

$$J = J_0 \sin\delta = J_0 \sin(\theta_2 - \theta_1) \qquad ...(10)$$

Here J_0 is proportional to T and represents the maximum zero voltage current that can be passed by the junction. With zero applied voltage a dc current will flow across the junction with a value between J_0 and $-J_0$ depending on the value of the phase difference $(\theta_2 - \theta_1)$. Current-voltage characteristic is shown in Fig. 44.2. The effect is called *dc Josephson effect.*

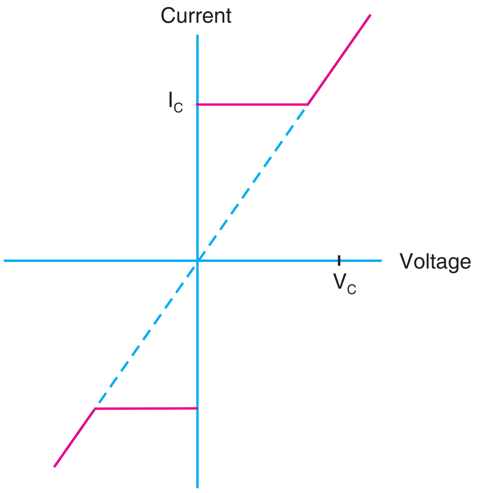


Fig. 44.2

44.3 Theory of AC Josephson Effect

Let a voltage V be applied across the Josephson junction. This is possible because the junction is an insulator. An electron pair experiences a potential energy difference qV on passing across the junction, where $q = (-2e)$. One can say that a pair on one side is at potential energy $-eV$ and the pair on the other side is at eV. The equation of motion of a pair are :

$$i\hbar \frac{\partial \psi_1}{\partial t} = \hbar T \psi_2 - eV\psi_1$$

$$i\hbar \frac{\partial \psi_2}{\partial t} = \hbar T \psi_1 + eV\psi_2 \qquad ...(1)$$

Let $\psi_1 = n_1^{1/2} e^{i\theta_1}$ and $\psi_2 = n_2^{1/2} e^{i\theta_2}$

$$\frac{\partial \psi_1}{\partial t} = \frac{1}{2} n_1^{-1/2} e^{i\theta_1} \frac{\partial n_1}{\partial t} + n_1^{1/2} e^{i\theta_1} \frac{\partial \theta_1}{\partial t}$$

$$i\hbar\, e^{i\theta_1} \left[\frac{n_1^{-1/2}}{2} \left(\frac{\partial n_1}{\partial t} \right) + (n_1^{1/2}\, i) \frac{\partial \theta_1}{\partial t} \right] = \hbar\, T\, n_2^{1/2} e^{i\theta_2} - eV\, n_1^{1/2} e^{i\theta_1}$$

Dividing throughout by $i\hbar \;\; e^{i\theta_1} n_1^{-\frac{1}{2}}$, we get

$$\frac{1}{2} \frac{\partial n_1}{\partial t} + i n_1 \frac{\partial \theta_1}{\partial t} = ie\, V n_1\, (\hbar)^{-1} - iT\, (n_1\, n_2)^{1/2}\, e^{i\delta} \qquad ...(2)$$

$$\frac{1}{2} \left(\frac{\partial n_1}{\partial t} \right) + i n_1 \left(\frac{\partial \theta_1}{\partial t} \right) = ien_1\, V\, (\hbar)^{-1} - iT\, [\cos\delta + i \sin\delta]\, (n_1\, n_2)^{1/2}$$

Breaking this into real and imaginary parts, we obtain

$$\frac{\partial n_1}{\partial t} = 2T\, (n_1\, n_2)^{1/2} \sin\delta \qquad ...(3)$$

$$\frac{\partial \theta_1}{\partial t} = \frac{eV}{\hbar} - T \left(\frac{n_2}{n_1} \right)^{1/2} \cos\delta \qquad ...(4)$$

This differs from Eq. (7) of Art. 44.2 by the term $\frac{eV}{\hbar}$. Further, by extension of Eq. (3) of Art. 44.2 we obtain

$$\frac{1}{2} \frac{\partial n_2}{\partial t} + i n_2 \frac{\partial \theta_2}{\partial t} = -ieVn_2\, \hbar^{-1} - iT\, (n_1\, n_2)^{1/2}\, e^{-i\delta} \qquad ...(5)$$

Here, $$\frac{\partial n_2}{\partial t} = -2T\, (n_1\, n_2)^{1/2} \sin\delta \qquad ...(6)$$

$$\frac{\partial \theta_2}{\partial t} = -\left(\frac{eV}{\hbar} \right) - T \left(\frac{n_1}{n_2} \right)^{1/2} \cos\delta \qquad ...(7)$$

Taking $n_1 \approx n_2$, we obtain from Eqs. (4) and (7)

$$\frac{\partial (\theta_2 - \theta_1)}{\partial t} = \frac{\partial \delta}{\partial t} = -\frac{2eV}{\hbar} \qquad ...(8)$$

Integrating Eq. (8), we see that with a *dc* voltage across the junction, the relative phase of probability amplitudes vary as

$$\delta(t) = \delta(0) - \frac{2eVt}{\hbar} \qquad ...(9)$$

The supercurrent is now given by

$$J = J_0 \sin\left[\delta(0) - \frac{2eVt}{\hbar}\right] \qquad ...(10)$$

This shows that the current oscillates with a frequency

$$\omega = \frac{2eV}{\hbar} \qquad ...(11)$$

This is called the *ac Josephson effect.* It permits a determination of $\frac{e}{\hbar}$ from the relationship between voltage and the frequency of the photon emitted when an electron pair crosses the barrier. A *dc* voltage of 1 μ *V* produces a frequency of 483.6 MHz.

Macroscopic Quantum Interference

The relation Eq. (10) provides the basis for one family of Josephson junction device called SQUID (superconducting quantum interference device). The SQUID (Fig. 44.3) is a double-junction quantum interoferometer formed from two Josephson junctions mounted on a superconducting ring. Magnetic field is applied normal to the plane of the ring.

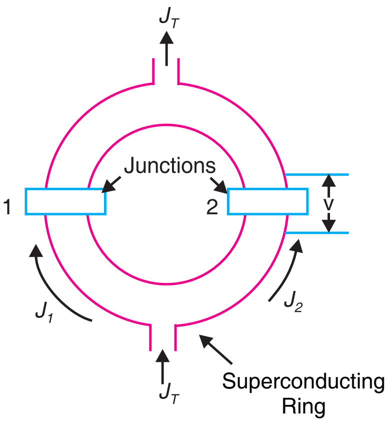


Fig. 44.3

In this device we may write the phases as

$$\delta_1 = \delta_0 - \frac{e}{\hbar c}\phi$$

and

$$\delta_2 = \delta_0 + \frac{e}{\hbar c}\phi \qquad ...(12)$$

where ϕ is the magnetic flux.

The total current, which is the sum of J_1 and J_2, is

$$J_{total} = J_0\left\{\sin\left(\delta_0 + \frac{e}{\hbar c}\phi\right) + \sin\left(\delta_0 - \frac{e}{\hbar c}\phi\right)\right\}$$

$$= 2(J_0 \sin\delta_0)\cos\frac{e}{\hbar\, c}\phi \qquad ...(13)$$

The total current varies with ϕ. The magnitude of the total current has maxima when

$$\frac{e}{\hbar c}\phi = s\pi,$$

where s is an integer ...(14)

The periodic variation of the critical current with magnetic flux is shown in Fig. 44.4.

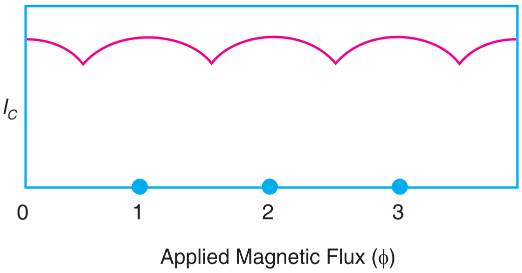


Fig. 44.4

44.4 Flux Quantization

The SQUID is a ring of superconductor interrupted by a Josephson junction. Let us now look at the phenomenon of *flux quantization.* Suppose we form a ring of superconducting material, place it in a magnetic field, and then reduce the temperature below T_c. While the Meissner effect excludes the field from the superconducting material itself, the field is still able to thread through the hole in the ring. If the external field is turned off, the flux through the hole is unable to decrease. The reason is Faraday's law. The rate of change of flux Φ through an area is the integral of the electric field E around the perimeter of the area:

$$\oint \vec{E}.\vec{dl} = -\frac{\partial \phi}{\partial t} \qquad ...(1)$$

If, as illustrated in Fig. 44.5 (*a*), we take the circuit of integration around the ring within the material, then, since E is zero inside any superconductor, we must have $\partial \Phi / \partial t = 0$. That is, the flux is trapped as Fig. 44.5 (*b*) indicates. It is maintained by the flow of surface currents in a thin layer of depth λ in the surface of the superconductor. An important feature here, that we have not derived, is that the amount of flux threading through the ring is *quantized.* It is known that

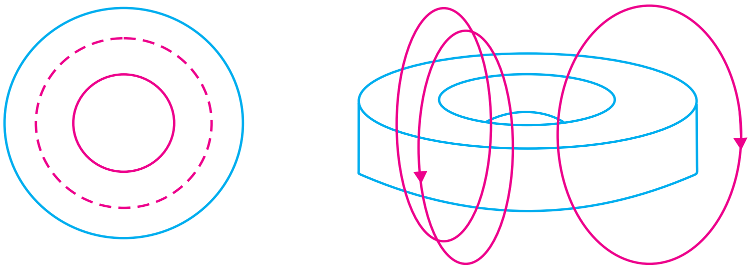


Fig. 44.5

$$\Phi = \frac{2\pi n\hbar}{q}$$

where n is an integer and q is the charge on the carrier of the supercurrent. Since the charge carriers are Cooper pairs, the value of q is $2e$. The result, that the minimum value of the quantized flux is

$$\Phi_0 = \frac{\pi\hbar}{e} \qquad ...(2)$$

and not $2\pi\hbar/e$, has been verified experimentally. London originally predicted this effect before the BCS theory and assumed $q = e$. The result $q = 2e$ is thus a strong element in the verification of the BCS theory.

The basic unit of quantized flux is

$$\phi_0 = \frac{h}{2e} = \frac{6.6256 \times 10^{-34}}{2(1.6021 \times 10^{-19})} = 2.068 \times 10^{-15} \text{ Wb.}$$

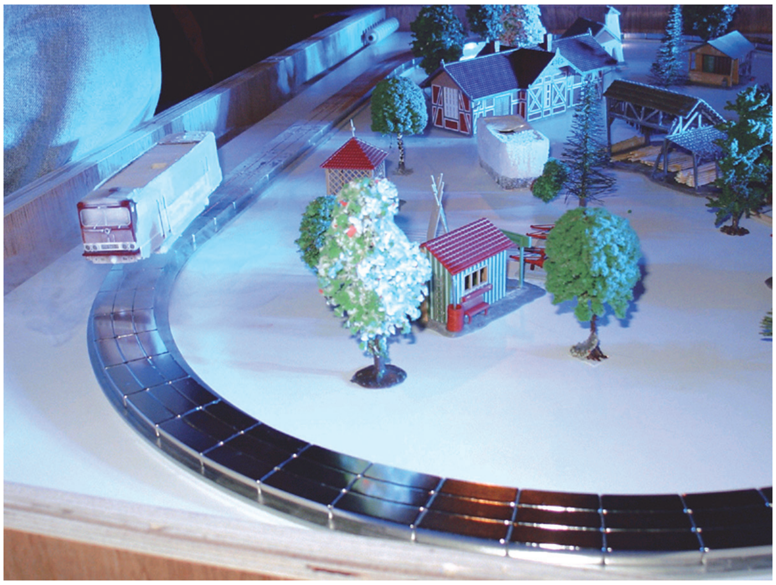
Quantized lines of flux can also be trapped within the interiors of certain forms of superconductors. The Meissner effect is not really violated here because the "hole" through which the field is threaded is a line of *normal material.* Such superconductors are designated as *type–II.* The critical fields in these materials are much higher than in the usual *type–I* form of superconductor, which does not allow such penetration by magnetic fields without destruction of the superconducting state.

44.5 High-T_c Superconductivity

A large number of possible practical applications of the phenomenon of superconductivity has not been achieved on account of the limit to the T_c value. Because at low T_c, the maintenance of superconductivity through refrigeration requires really a very high cost. In last couple of years, special efforts have been made by researchers to discover high T_c superconductors. The corresponding data are displayed in Table 44.1.

Table 44.1: High-T_c Superconductors

	Material	T_c
1.	Cu-O Perovskite-type	30 to 100 K
2.	$La_{1.85}\,Sr_{0.15}\,CuO_4$	36 K
3.	$La_{1.8}\,Sr_{0.2}\,CuO_4$	40 K
4.	$(Y_{0.4}\,Br_{0.6})\,CuO_3$	77 K
5.	$Y\,Ba_2\,Cu_3\,O_7$	90 K
6.	Granular Y — Ba —Cu—O	160 K
7.	Lanthanum compounds	300 to 500 K



A pellet of Superconductor with High Critical Temperature.

The detailed mechanism for superconductivity in the high-T_c materials is not at present known precisely. A number of arguments have been advanced to explain high T_c. For example, it has been realised that the role of oxygen is essential for high T_c oxide superconductors. The occurrence of a strong electronic band in the infrared spectrum and its correlation with superconductivity provides evidence that the superconductivity of $Y\,Ba_2\,Cu_3\,O_{7-y}$ materials is mediated by an excitonic mechanism. Pauling has endeavoured to support Anderson's resonating valence bond theory of superconductivity in compounds such as $(Sr, Ba, Y, La)_2\,CuO_{4-y}$: the conductivity is increased by the resonance along the O—Cu ... O—Cu lines of atoms, and superconductivity is probably achieved through interaction with a phonon.

44.6 Applications

There are many applications of superconductivity.

(*i*) One is in the construction of superconducting electromagnets that carry large resistanceless currents and, therefore, produce large magnetic fields (≈ 5 – 10 Tesla). If we use superconducting wire for magnet windings, currents of the order of 100 amperes can be carried by very fine wires. Thus small-size magnets can be constructed. The principle is illustrated in Fig. 44.6, in which a superconducting coil is immersed in liquid helium (4 K). Starting from zero, the current is increased until the magnetic field B reaches the desired value. At that point the switch S is closed. The current now flows through the switch. The power supply is then turned off. The current will continue to flow through the coil indefinitely without resistive losses.

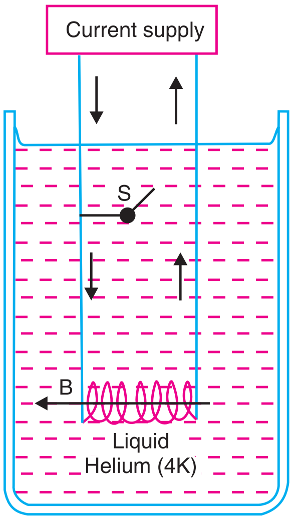
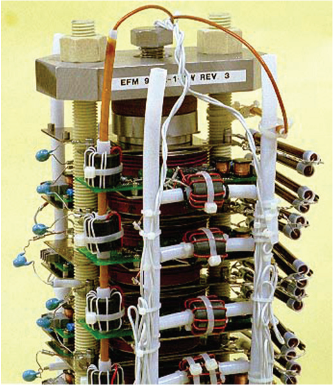


Fig. 44.6

(*ii*) Superconducting cables can be used to transmit electric power over long distances without resistive losses. This is possible only if the cost of keeping the cable below its critical temperature is less than the value of the power lost using ordinary cables.



Switching Devices

(*iii*) Superconductors are used for amplifying very small direct currents and voltages.

(*iv*) Superconductors are employed in switching devices.

(*v*) Because superconductors are diamagnetic, they can be used to shield out unwanted magnetic flux, as in shaping the magnetic lens system of an electron microscope.

(*vi*) The most important electronic device based on superconductors is the SQUID (superconducting quantum interference device). It consists of two **Josephson junctions** (*i.e.,* two superconductors separated by an insulator). The current flowing through such a junction is a function of the voltage or magnetic field across it. Using such a device, very weak magnetic fields can be detected. Microwave and far infrared **radiation detectors** and other microwave applications like **wave guides, mixers,** etc., are fabricated using high T_c superconductors.

EXERCISE

1. Explain *dc Josephson effect.* Show that the supercurrent of superconducting pairs across the junction depends on the phase difference.
2. Discuss *ac Josephson effect.* Show that the current oscillates with frequency $\omega = 2eV/\hbar$.
3. Discuss *Josephson effect.* Give formulation of *ac Josephson effect.* Explain how *ac Josephson effect* can estimate $e/\hbar$. (*Raj.,* 1987)
4. Give the principle of SQUID.
5. What is the frequency of the electromagnetic waves radiated by a Josephson junction having a voltage of 650 μV across its terminals ? (**Ans.** 3.15×10^{11} Hz)

CHAPTER

45

SUPERCONDUCTIVITY—II

AT A GLANCE			
45.1	London Equation	45.2	Energy Gap
45.3	Microwave and Infrared Absorptions		

45.1 London Equation

Electrical conduction in the normal state of a metal is described by Ohm's law $\mathbf{j} = \sigma\mathbf{E}$. We modify this to describe conduction and the Meissner effect in the superconducting state. We assume that in the superconducting state the current density is directly proportional to the vector potential **A** of the local magnetic field. Here, $\mathbf{B} = \text{curl } \mathbf{A}$. The constant of proportionality is $-1/\mu_0 \lambda_L^2$. Thus

$$\mathbf{j} = -\frac{1}{\mu_0 \lambda_L^2}\mathbf{A} \qquad \text{...(1)}$$

Here, λ_L is a constant with the dimensions of length.

This is the London equation.

Taking the curl of both sides, we get

$$\text{curl } \mathbf{J} = -\frac{1}{\mu_0 \lambda_L^2}\mathbf{B} \qquad \text{...(2)}$$

Now, consider the Maxwell equation

$$\text{curl } \mathbf{B} = \mu_0 \mathbf{j} \qquad \text{...(3)}$$

Taking curl, we obtain

$$\text{curl curl } \mathbf{B} = \mu_0 \text{ curl } \mathbf{j} \qquad \text{...(4)}$$

Now, $\text{curl curl } \mathbf{B} = \text{grad div } \mathbf{B} - \nabla^2\mathbf{B}$

$\text{curl curl } \mathbf{B} = -\nabla^2\mathbf{B} \quad (\because \text{div } \mathbf{B} = 0)$ ...(5)

On combining Eqs. (4) and (5),

$$-\nabla^2\mathbf{B} = \mu_0 \,\text{curl}\, \mathbf{j}$$

Using London Eq. (2), we get

$$-\nabla^2\mathbf{B} = \mu_0\left(-\frac{1}{\mu_0\lambda_L^2}\mathbf{B}\right)$$

$$\therefore \qquad -\nabla^2\mathbf{B} = \mathbf{B}/\lambda_L^2 \qquad \text{... (6)}$$

Here, λ_L is called the *London penetration depth.*

For a superconductor to the right of the plane $x = 0$, Eq. (6) has the solution

$$B(x) = B(0)\exp(-x/\lambda_L), \qquad \text{... (7)}$$

Here, B (0) is the field at the plane boundary.

The penetration depth λ_L is defined as the distance in which the field decreases by the factor e^{-1}.

Eq. (7) indicates that **B** does not penetrate very deeply into a superconductor. Therefore, it implies the *Meissner effect.*

A magnetic field external to the superconductor decays exponentially as it penetrates the superconductor (Fig. 45.1). The decay constant is called the penetration depth λ_L. The vacuum-superconducting interface is at $x = 0$. Eq. (7) agrees well with the experimental observations that the magnetic field does not drop to zero abruptly at the surface but decays inside the superconductor in a characteristic distance λ_L, called *London penetration depth.* Thus the flux vanishes in the interior of the superconductor. The values of λ_L found experimentally range from 30 nm to 500 nm.

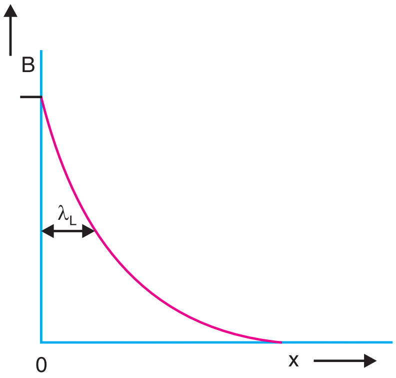


Fig. 45.1

45.2 Energy Gap

Experiments have shown that in superconductors, for temperatures in the vicinity of absolute zero, a forbidden energy gap just above the Fermi level is observed. Figure 45.2 (*a*) shows the conduction band in the normal state. Fig. 45.2 (*b*) depicts an energy gap equal to $2E_g$ at the Fermi level in the superconducting state. Thus, the Fermi level in a superconductor is midway between the ground state and the first excited state so that each lies an energy distance = E_g away from the Fermi level.

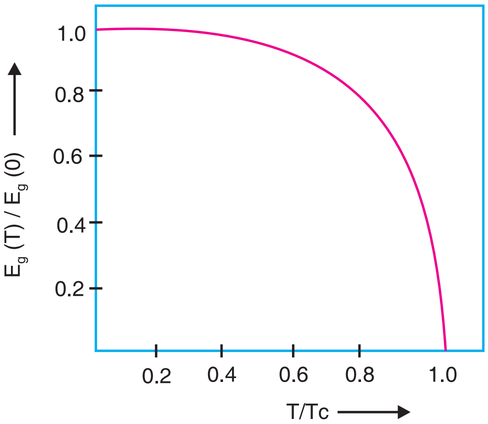


Fig. 45.3

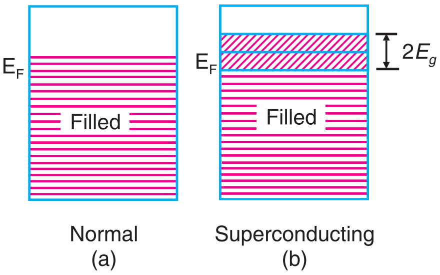


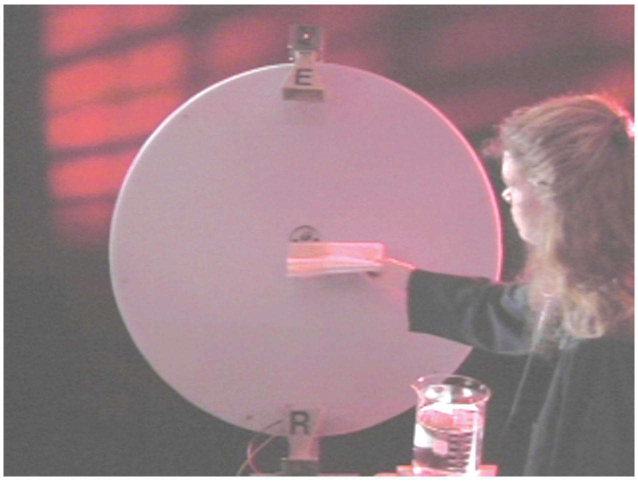
Fig. 45.2

Electrons in excited states above the gap behave as normal electrons. At absolute zero, there are no electrons above the gap. E_g is typically of the order of 10^{-4} eV.

E_g is found to be a function of temperature T. Figure 45.3 shows reduced values of observed energy gap $E_g(T) / E_g(0)$ as a function of the reduced temperature T/T_c . We observe that the energy gap decreases continuously to zero as the temperature is increased to T_c . The transition from the superconducting state to the normal state is observed to be a *second-order phase transition*. In such a transition, there is no latent heat, but there is a discontinuity in the heat capacity.

45.3 Microwave and Infrared Absorptions

The response of a metal to electromagnetic radiation is determined by the frequency dependent conductivity. This in turn depends on the available mechanisms for energy absorption by the conduction electrons at the given frequency. The electronic excitation spectrum in the superconducting state is characterised by an energy gap $2E_g$. So we expect the *AC* conductivity to differ substantially from its normal state form at frequencies small compared with $2E_g/\hbar$ and to be essentially the same in the superconducting and normal states at frequencies large compared with $2E_g/\hbar$, The value of $2E_g/\hbar$ is typically in the range between microwave and infrared frequencies. In the superconducting state, an *AC* behaviour is observed which is indistinguishable from that in the normal state at optical frequencies. Deviations from normal state behaviour first appear in the infrared. At microwave frequencies *AC* behaviour fully displaying the lack of electronic absorption characteristic of an energy gap becomes completely developed.



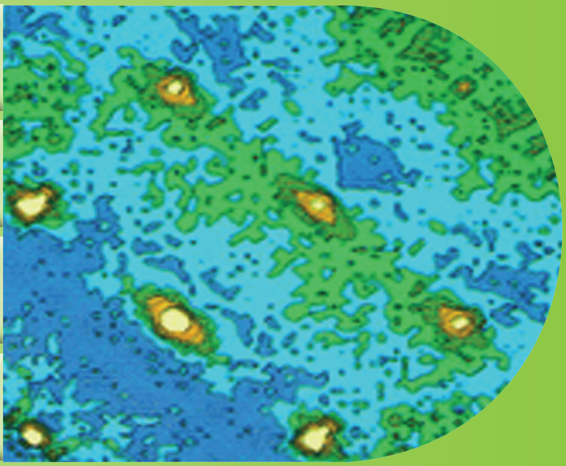
Microwave Absorption.

EXERCISE

1. What is Meissner effect? Obtain an expression for the London penetration depth of magnetic field for a superconductor. (*Raj.*, 1976)

CHAPTER

46



RECIPROCAL LATTICE, CRYSTAL BINDING, ENERGY BANDS

AT A GLANCE

46.1	Reciprocal Lattice	**46.2**	Reciprocal Lattice Vectors
46.3	Properties of Reciprocal Lattice	**46.4**	Bragg's Law for Reciprocal Lattice
46.5	Brillouin Zones		
46.6	Reciprocal Lattice to Simple Cubic (sc) Lattice		
46.7	Reciprocal Lattice to bcc Lattice	**46.8**	Reciprocal Lattice to fcc Lattice
46.9	Atomic Scattering Factor	**46.10**	The Geometrical Structure Factor
46.11	Cohesive Energy of Inert Gas Crystals	**46.12**	Binding Energy of Ionic Crystals
46.13	Evaluation of the Madelung Constant	**46.14**	Bloch Theorem
46.15	Kronig-Penney Model	**46.16**	Effective Mass of Electron
46.17	Interatomic Potential		
46.18	Second Order Expansion of Potential Energy		
46.19	Second Order Expansion of Lennard-Jones Potential about the Minimum		
46.20	Simple Harmonic Oscillator		

RECIPROCAL LATTICE

46.1 Reciprocal Lattice

In a crystal lattice there exist many sets of planes, with different orientations and spacings, which can cause diffraction. Fig. 46.1 shows three sets of lattice planes with interplanar spacings d_1, d_2 and d_3. Let us draw, from a common origin, normals to all sets of planes, the length of each normal being proportional to the reciprocal of the interplanar spacing of the corresponding set. Then, the end-points of the normals form a lattice which is called the '*reciprocal lattice*'.

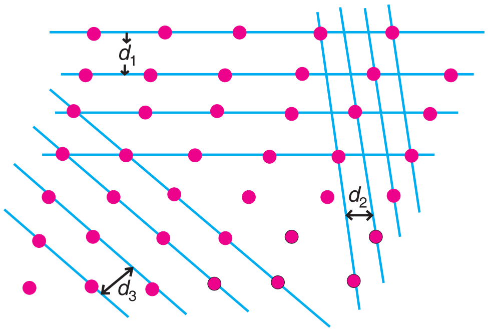


Fig. 46.1

Each point in the reciprocal lattice preserves the characteristics of the set of planes which it represents. Its direction with respect to the origin represents the orientation of the planes, and its distance from the origin represents the interplanar spacing of the planes.

46.2 Reciprocal Lattice Vectors

Let $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$ be the primitive vectors of the (direct) crystal lattice. Then, primitive vectors $\mathbf{b}_1$, $\mathbf{b}_2$, $\mathbf{b}_3$ of the reciprocal lattice are defined by

$$\mathbf{b}_1 = 2\pi\frac{\mathbf{a}_2\times\mathbf{a}_3}{\mathbf{a}_1\cdot\mathbf{a}_2\times\mathbf{a}_3};\ \mathbf{b}_2 = 2\pi\frac{\mathbf{a}_3\times\mathbf{a}_1}{\mathbf{a}_1\cdot\mathbf{a}_2\times\mathbf{a}_3};\ \mathbf{b}_3 = 2\pi\frac{\mathbf{a}_1\times\mathbf{a}_2}{\mathbf{a}_1\cdot\mathbf{a}_2\times\mathbf{a}_3}\cdot \quad ...(1)$$

Each vector defined by Eq. (1) is orthogonal to two axis vectors of the crystal lattice. Thus $\mathbf{b}_1$, $\mathbf{b}_2$, $\mathbf{b}_3$ have the property

$$\mathbf{b}_i \cdot \mathbf{a}_j = 2\pi\delta_{ij}, \quad ...(2)$$

Here, $\delta_{ij} = 1$ if $i = j$ and $\delta_{ij} = 0$ if $i \neq j$.

Points in the reciprocal lattice are mapped by the set of vectors

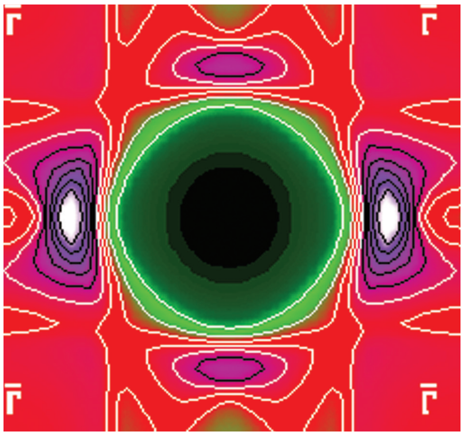
$$\mathbf{G} = \upsilon_1\mathbf{b}_1 + \upsilon_2\mathbf{b}_2 + \upsilon_3\mathbf{b}_3. \quad ...(3)$$

Here, υ_1, υ_2, υ_3 are integers. A vector **G** of this form is a *reciprocal lattice vector*.

It connects any two points of the reciprocal lattice.

46.3 Properties of Reciprocal Lattice

1. Every crystal structure has two lattices associated with it, the crystal lattice and the reciprocal lattice. A diffraction pattern of a crystal is a map of the reciprocal lattice of the crystal. A microscope image is a map of the crystal structure in real space. Thus when we rotate a crystal in a holder, we rotate both the direct lattice and the reciprocal lattice.
2. Vectors in the direct lattice have the dimensions of [length]. Vectors in the reciprocal lattice have the dimensions of length^{-1}.
3. A crystal lattice is a lattice in real space. The reciprocal lattice is a lattice in the Fourier space associated with the crystal.
4. Each point in a reciprocal lattice corresponds to particular set of parallel planes of the direct lattice.



One Cell of the Reciprocal Lattice

5. The distance of a reciprocal lattice point from an arbitrarily fixed origin is inversely proportional to the interplanar spacing of the corresponding parallel planes of the direct lattice.
6. The volume of a unit cell of the reciprocal lattice is inversely proportional to the volume of a unit cell of the direct lattice.
7. The unit cell of the reciprocal lattice need not be a parallelopiped. In fact, we almost always deal with the Wigner-Seitz cell of the reciprocal lattice. This is called a *Brillouin zone*.

8. The direct lattice is the reciprocal of its own reciprocal lattice.
9. Simple cubic lattice is its own reciprocal. bcc and fcc lattices are reciprocal to each other.

Importance. A diffraction pattern of a crystal is a map of the reciprocal lattice of the crystal. The Bragg's diffraction law can be expressed in terms of the reciprocal lattice vector **G**, as

$$2\,\mathbf{k}\,.\,\mathbf{G} + G^2 = 0.$$

Here, **k** is the wave vector of the incident wave. This equation is very important in wave propagation and elastic scattering in crystal lattices. A similar result arises in the theory of the electron energy band structure of crystals. Thus the concept of reciprocal lattice is very important in the study of lattice dynamics of solids.

46.4 Bragg's Law for Reciprocal Lattice

The Ewald construction for diffraction of X-rays is shown in Fig. 46.2. Draw a vector **AO** in the incident direction of length 1/λ terminating at the origin *O*.

AE = Reflecting plane

$$AO = AB = \frac{1}{\lambda}$$

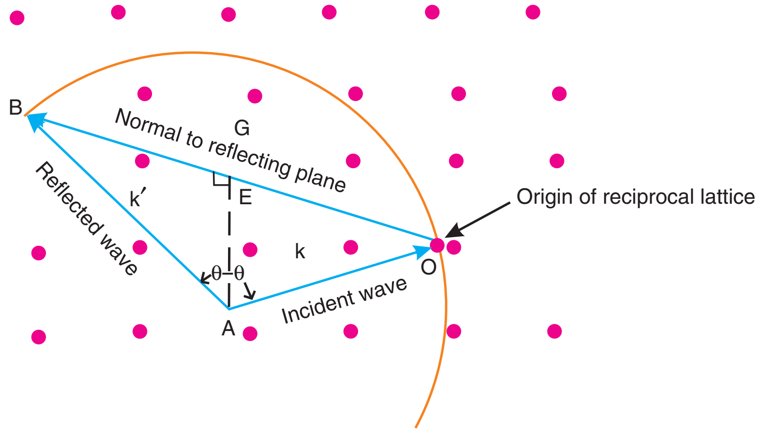


Fig. 46.2

Construct a circle of radius 1/λ (a sphere, called *reflex sphere*, of radius 1/λ in three dimensions) with centre at *A*. The circle passes through any point *B* of the reciprocal lattice. Join *A* and *O* to *B*. Thus, **OB** is a reciprocal lattice vector, **G. OB** is normal to the lattice plane *AE*.

Hence, $OB = |\mathbf{G}| = 1/d$, where *d* is the interplanar spacing. The wave vectors of the incoming and outgoing beams are **k** = **AO** and **k'** = **AB**. Thus,

$$\mathbf{k}' = \mathbf{k} + \mathbf{G} \qquad ...(1)$$

For diffraction, it is necessary that the vector **k**' *be equal in magnitude to the vector* **k**:

$$\mathbf{k}'^2 = (\mathbf{k} + \mathbf{G})^2 = \mathbf{k}^2$$

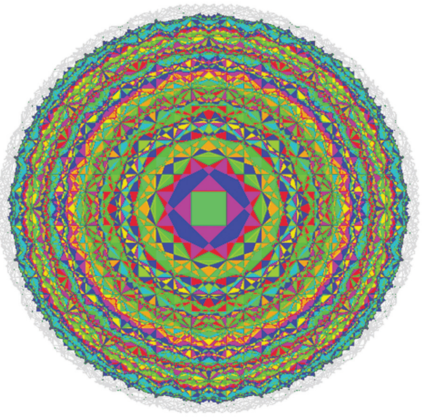
or $$2\mathbf{k}\,.\,\mathbf{G} + \mathbf{G}^2 = 0 \qquad ...(2)$$

This is the Bragg's law in vector form.

46.5 Brillouin Zones

A Brillouin zone is defined as a Wigner-Seitz primitive cell in the 'reciprocal' lattice.

The first Brillouin zone of an oblique lattice in two dimensions is constructed in Fig. 46.3. We first draw an adequate number of vectors from O to nearby points in the reciprocal lattice. Next we draw lines perpendicular to these vectors at their midpoints. The *smallest* enclosed area is the first Brillouin zone. Any wave whose wave vector drawn from the origin terminates on the boundary of the zone will be diffracted by the crystal. *The first Brillouin zone is thus the smallest volume entirely enclosed by planes which are the perpendicular bisectors of the reciprocal lattice vectors drawn from the origin.*



Two-Dimensional Coloured Brillouin Zones

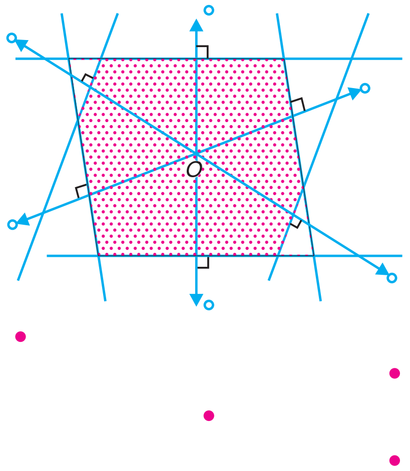


Fig. 46.3

46.6 Reciprocal Lattice to Simple Cubic (sc) Lattice

The primitive translation vectors of a simple cubic lattice may be written as

$$\mathbf{a}_1 = a\hat{\mathbf{x}}; \qquad \mathbf{a}_2 = a\hat{\mathbf{y}}; \qquad \mathbf{a}_3 = a\hat{\mathbf{z}} \qquad \text{...(1)}$$

Here $\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}$ are orthogonal vectors of unit length.

The volume of the unit cell

$$= \mathbf{a}_1 \cdot \mathbf{a}_2 \times \mathbf{a}_3 = a^3. \qquad \text{...(2)}$$

The primitive translation vectors of the reciprocal lattice are

$$\mathbf{b}_1 = (2\pi/a)\,\hat{\mathbf{x}}; \qquad \mathbf{b}_2 = (2\pi/a)\,\hat{\mathbf{y}}; \quad \mathbf{b}_3 = (2\pi/a)\,\hat{\mathbf{z}} \qquad \text{...(3)}$$

Eqs. (3) indicates that all the three reciprocal *lattice* vectors are equal in magnitude. Thus the reciprocal lattice is itself a simple cubic lattice, now of lattice constant $2\pi / a$ (Fig. 46.4).

The boundaries of the first Brillouin zones are the planes normal to the six reciprocal lattice vectors $\pm\,\mathbf{b}_1, \pm\,\mathbf{b}_2, \pm\,\mathbf{b}_3$, at their midpoints.

$$\pm\frac{1}{2}\mathbf{b}_1 = \pm(\pi/a)\hat{\mathbf{x}}; \quad \pm\frac{1}{2}\mathbf{b}_2 = \pm(\pi/a)\hat{\mathbf{y}}; \quad \pm\frac{1}{2}\mathbf{b}_3 = \pm(\pi/a)\hat{\mathbf{z}}.$$

The six planes bound a cube of edge $2\pi / a$ and of volume $(2\pi / a)^3$. This cube is the first Brillouin zone of the sc crystal lattice.

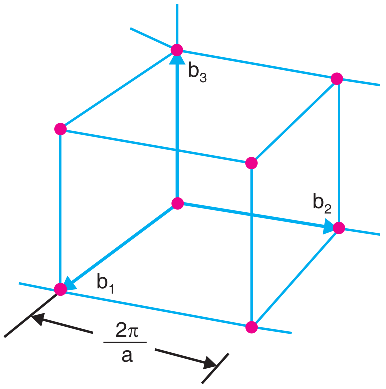


Fig. 46.4

46.7 Reciprocal Lattice to bcc Lattice

The primitive translation vectors of the bcc lattice (Fig. 46.5) are

$$\mathbf{a}_1 = \frac{1}{2}a(-\hat{\mathbf{x}}+\hat{\mathbf{y}}+\hat{\mathbf{z}}); \quad \mathbf{a}_2 = \frac{1}{2}a(\hat{\mathbf{x}}-\hat{\mathbf{y}}+\hat{\mathbf{z}}); \quad \mathbf{a}_3 = \frac{1}{2}a(\hat{\mathbf{x}}+\hat{\mathbf{y}}-\hat{\mathbf{z}}) \qquad ...(1)$$

Here, a is the side of the cube edge.

$\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}$ are orthogonal unit vectors parallel to the cube edges.

The volume of the primitive cell is

$$V = |\mathbf{a}_1 \cdot \mathbf{a}_2 \times \mathbf{a}_3| = \frac{1}{2}a^3 \qquad ...(2)$$

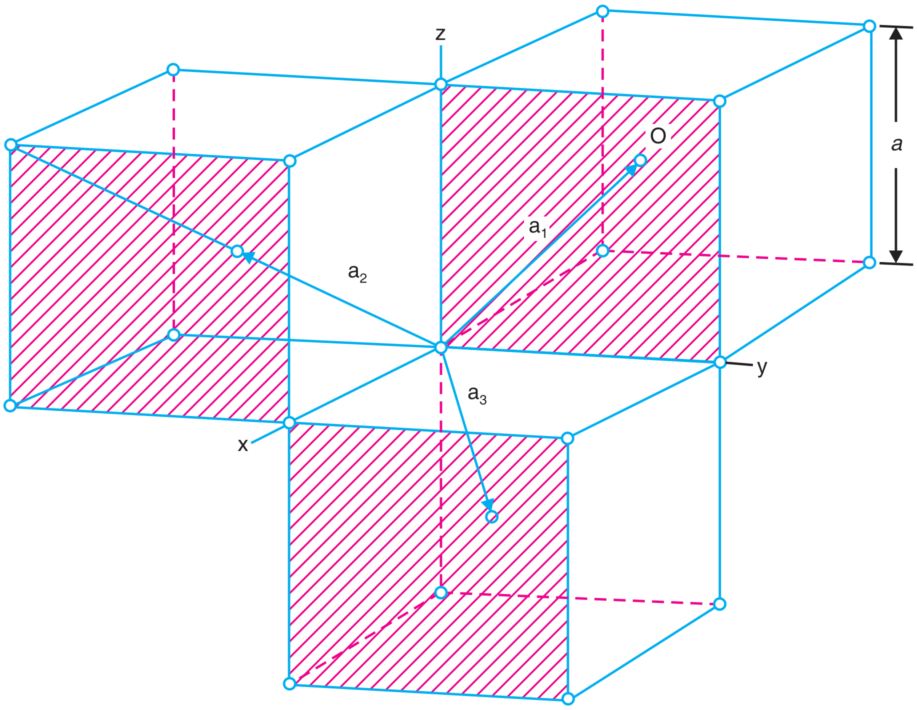


Fig. 46.5

The primitive translation vectors of the reciprocal lattice are

$$\mathbf{b}_1 = (2\pi/a)(\hat{\mathbf{y}}+\hat{\mathbf{z}}); \quad \mathbf{b}_2 = (2\pi/a)(\hat{\mathbf{x}}+\hat{\mathbf{z}}); \quad \mathbf{b}_3 = (2\pi/a)(\hat{\mathbf{x}}+\hat{\mathbf{y}}).$$

These are the primitive translation vectors of an fcc lattice.

Thus the reciprocal lattice to a bcc lattice is an fcc lattice.

Let υ_1, υ_2, υ_3, be integers.

The general reciprocal lattice vector is,

$$\mathbf{G} = \upsilon_1\mathbf{b}_1 + \upsilon_2\mathbf{b}_2 + \upsilon_3\mathbf{b}_3 = (2\pi/a)[(\upsilon_2+\upsilon_3)\hat{\mathbf{x}} + (\upsilon_1+\upsilon_3)\hat{\mathbf{y}} + (\upsilon_1+\upsilon_2)\hat{\mathbf{z}}] \qquad ...(4)$$

The shortest **G**'s are the following 12 vectors, where all choices of sign are independent:

$$(2\pi/a)(\pm\,\hat{\mathbf{y}} \pm \hat{\mathbf{z}}); \quad (2\pi/a)(\pm\,\hat{\mathbf{x}} \pm \hat{\mathbf{z}}); \quad (2\pi/a)(\pm\,\hat{\mathbf{x}} \pm \hat{\mathbf{y}}). \qquad ...(5)$$

The first Brillouin zone boundaries are determined by the twelve planes normal to these vectors at their midpoints.

The zone thus formed is a regular 12-faced solid, a rhombic dodecahedron (Fig. 46.6).

The vectors from the origin to the centre of each face of the first *BZ* are

$$(\pi/a)\ (\pm\hat{\mathbf{y}}\pm\hat{\mathbf{z}});\quad (\pi/a)\ (\pm\hat{\mathbf{x}}\pm\hat{\mathbf{z}});$$

$$(\pi/a)\ (\pm\hat{\mathbf{x}}\pm\hat{\mathbf{y}})\,. \qquad \text{...(6)}$$

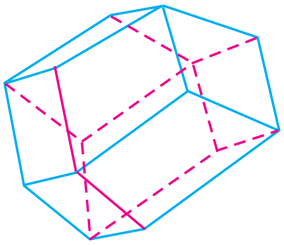


Fig. 46.6

46.8 Reciprocal Lattice to fcc Lattice

The primitive translation vectors of the fcc lattice (Fig. 46.7) are

$$\mathbf{a}_1 = \frac{1}{2}a\,(\hat{\mathbf{y}}+\hat{\mathbf{z}});\quad \mathbf{a}_2 = \frac{1}{2}a(\hat{\mathbf{x}}+\hat{\mathbf{z}});$$

$$\mathbf{a}_3 = \frac{1}{2}a(\hat{\mathbf{x}}+\hat{\mathbf{y}}). \qquad \text{...(1)}$$

The volume of the primitive cell is

$$V = |\mathbf{a}_1\cdot\mathbf{a}_2\times\mathbf{a}_3| = \frac{1}{4}a^3 \qquad \text{...(2)}$$

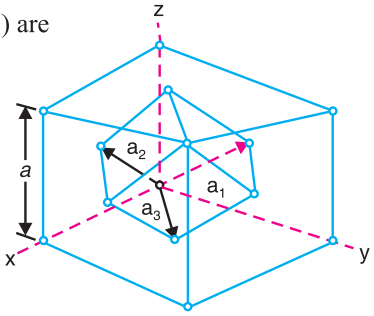


Fig. 46.7

The primitive translation vectors of the reciprocal lattice are

$$\mathbf{b}_1 = (2\pi/a)(-\hat{\mathbf{x}}+\hat{\mathbf{y}}+\hat{\mathbf{z}});$$

$$\mathbf{b}_2 = (2\pi/a)(\hat{\mathbf{x}}-\hat{\mathbf{y}}+\hat{\mathbf{z}});$$

$$\mathbf{b}_3 = (2\pi/a)(\hat{\mathbf{x}}+\hat{\mathbf{y}}-\hat{\mathbf{z}}). \qquad \text{...(3)}$$

These are primitive translation vectors of a bcc lattice having length of the cube edge as $(2\pi / a)$. Thus the reciprocal lattice to an fcc lattice is a bcc lattice.

The shortest **G**'s are the following eight vectors

$$(2\pi/a)(\pm\hat{\mathbf{x}}\pm\hat{\mathbf{y}}\pm\hat{\mathbf{z}}) \qquad \text{...(4)}$$

The boundaries of the first Brillouin zone are determined mostly by the normal bisector planes to the above eight vectors. But the corners of the octahedron thus formed are cut by the planes that are the perpendicular bisectors of six other reciprocal lattice vectors;

$$(2\pi/a)(\pm 2\hat{\mathbf{x}});\qquad (2\pi/a)(\pm 2\hat{\mathbf{y}});$$

$$(2\pi/a)(\pm 2\hat{\mathbf{z}}). \qquad \text{...(5)}$$

The first Brillouin zone has the shape of the truncated octahedron (Fig. 46.8).

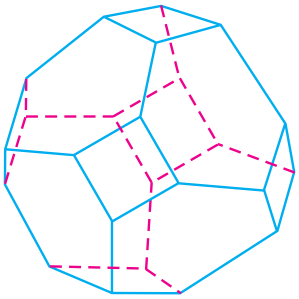


Fig. 46.8

46.9 Atomic Scattering Factor

The atomic scattering factor, f, describes the scattering power of a single atom in relation to the scattering power of a single electron.

$$f = \frac{\text{amplitude of the radiation scattered from the atom}}{\text{amplitude of the radiation scattered from an electron}}$$

This factor 'f' is called the atomic scattering factor or form factor.

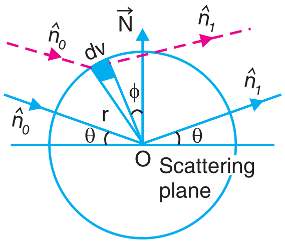


Fig. 46.9

Calculation of *f*

Consider an atom containing electrons arranged in a spherically symmetric configuration around its centre which is taken as origin *O* (Fig. 46.9). Let r be the radius of the atom. Let $\rho(\mathbf{r})$ be the charge density at a point $\mathbf{r}$. Consider a small volume element dv at $\mathbf{r}$.

Charge located at $\mathbf{r} = \rho(\mathbf{r})\, dv$.

Let $\hat{\mathbf{n}}_0$ and $\hat{\mathbf{n}}_1$ represent the directions of the incident and scattered beam respectively. We first consider the scattering from the charge $\rho(\mathbf{r})\, dv$ and an electron located at the origin.

The phase difference between the radiation scattered from the charge $\rho(\mathbf{r})\, dv$ and that scattered from the electron is,

$$\phi_r = \frac{2\pi}{\lambda}(\mathbf{r}.\mathbf{N}) \qquad \text{...(1)}$$

Here, $\mathbf{N}$ is the scattering normal.

Now, consider that the scattering amplitude from the point electron in the direction $\hat{\mathbf{n}}_1$ is represented by $Ae^{i(kx-\omega t)}$. Here, x is the distance coordinate along $\hat{\mathbf{n}}_1$ and k is the wave number. Then the scattering amplitude from the charge $\rho(\mathbf{r})\, dv$ in the same direction will be proportional to the magnitude of the charge and will contain the phase factor $e^{i\phi_r}$. It is represented by

$$Ae^{i(kx-\omega t)+i\phi_r}\rho(\mathbf{r})\,dv.$$

Therefore, the ratio of the amplitude of the radiation scattered by the charge element dv to that scattered by a point electron at the origin is,

$$df = \frac{Ae^{i(kx-\omega t)+i\phi_r}\rho(\mathbf{r})\,dv}{Ae^{i(kx-\omega t)}} = e^{i\phi_r}\rho(\mathbf{r})\,dv$$

Therefore, '*f*' the ratio of the scattered amplitude from the whole atom to that from a point electron is

$$f = \int_V \rho(\mathbf{r})e^{i\phi_r}\,dv \qquad \text{...(2)}$$

Here, V is the volume of the atom.

Electrons in the atom have spherically symmetric charge distribution. So $\rho(\mathbf{r})$ is a function of r only. The volume element dv in the spherical coordinate system is,

$$dv = 2\pi r^2 \sin\phi\, d\phi\, dr \qquad \text{...(3)}$$

We have, $|\mathbf{N}| = 2\sin\theta$

Therefore, Eq. (1) becomes

$$\phi_r = (2\pi/\lambda)\, rN\cos\phi = (4\pi/\lambda)\, r\sin\theta\cos\phi = \mu\, r\cos\phi \qquad \text{...(4)}$$

Here, $\mu = (4\pi/\lambda)\sin\theta$ ...(5)

Substitution of Eqs. (3) and (4) in Eq. (2) gives,

$$f = \int_{r=0}^{\infty}\int_{\phi=0}^{\pi} \rho(\mathbf{r})e^{i\mu r\cos\phi}\, 2\pi r^2 \sin\phi\, d\phi\, dr$$

Now $$\int_0^{\pi} e^{i\mu r\cos\phi}\sin\phi\, d\phi = \frac{2(\sin\mu r)}{\mu r}$$

$$\therefore \qquad f = \int_0^{\infty} 4\pi r^2 \rho(\mathbf{r})\frac{\sin\mu r}{\mu r}\,dr \qquad \text{...(6)}$$

This is the general expression for atomic scattering factor.

As $\theta \to 0$, $\mu \to 0$ and $(\sin \mu r) / \mu r \to 1$, we have

$$\lim_{\mu \to 0} f(\mu) = \int_0^{\infty} 4\pi r^2 \rho(r)\, dr = Z \cdot \qquad ...(7)$$

46.10 The Geometrical Structure Factor

The intensity of an x-ray beam diffracted by a crystal depends upon the atomic scattering factors of the various atoms involved. It depends also on the contents of the unit cell, *i.e.*, on the number, type and distribution of atoms in the unit cell. So if we wish to discuss the characteristics of the intensity of a particular order reflection and the relative intensities of the various reflections, we must take into account the contributions made by all the atoms in the unit cell to the scattered amplitude in a given direction.

Let us consider the $(h'k'l')$ reflection. The total scattering amplitude $F(h'k'l')$ for the reflection $(h'k'l')$ is defined as the ratio of the amplitude of radiation scattered by the entire unit cell to the amplitude of radiation scattered by a single point electron placed at origin for the same wavelength. It is given by

$$F(h'k'l') = \sum_j f_j e^{i\phi_j} = \sum_j f_j e^{i(2\pi/\lambda)(\mathbf{r}_j . \mathbf{N})} \qquad ...(1)$$

The summation extends over all the atoms belonging to a unit cell.

Here, f_j = the atomic scattering factor for the jth atom,

ϕ_j = phase difference between the radiation scattered from the jth atom of the unit cell and that scattered from the electron placed at the origin,

$\mathbf{r}_j$ = vector position of the **j**th atom,

$\mathbf{N}$ = scattering normal.

Let (u_j, v_j, w_j) represent the coordinates of jth atom. Then,

$$\mathbf{r}_j = u_j \mathbf{a} + v_j \mathbf{b} + w_j \mathbf{c} \qquad ...(2)$$

$$\therefore \quad \mathbf{r}_j . \mathbf{N} = \lambda (u_j h' + v_j k' + w_j l') \qquad ...(3)$$

Substituting Eq. (3) in Eq. (1), we get

$$F(h'k'l') = \sum_j f_j e^{i 2\pi (u_j h' + v_j k' + w_j l')} \qquad ...(4)$$

For identical atoms, all the f_j's have the same value f. Therefore, Eq. (4) becomes

$$F(h'k'l') = f\mathbf{S} \qquad ...(5)$$

Here, $$\mathbf{S} = \sum_j e^{2\pi i (u_j h' + v_j k' + w_j l')} \qquad ...(6)$$

S is called the *geometrical structure factor*. **S** depends upon the geometrical arrangement of atoms within the unit cell.

The intensity of the diffracted beam is proportional to the square of the amplitude F. $I = F^2 = F^* F$

$$= \left[\sum_j f_j \cos 2\pi (u_j h' + v_j k' + w_j l') \right]^2 + \left[\sum_j f_j \sin 2\pi (u_j h' + v_j k' + w_j l') \right]^2 \qquad ...(7)$$

Here F^* is the complex conjugate of F.

The scattered intensity vanishes in all directions except those in which the structure factor S is nonvanishing. These latter directions are therefore the directions of diffraction. When the Bragg condition is satisfied, then the incident beam is diffracted into a single beam, which is recorded at the detector as a single spot on a film. This spot represents the whole set of reflecting planes $(h'\,k'\,l')$. When the crystal is rotated so that a new set of planes again satisfies the Bragg condition, then this new set appears as a new spot on the film at the detector. Therefore each spot on the film represents a whole set of crystalline planes. From the arrangement of these spots one can determine the structure of the crystal.

Consider the bcc lattice. The unit cell has two atoms whose coordinates are

$(u_j, v_j, w_j) = (0,0,0)$ and $\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)$.

From Eq. (4), $$F(h'\,k'\,l') = f[1 + e^{i\pi(h'+k'+l')}]. \qquad ...(8)$$

This expression can take only two values. When $(h' + k' + l')$ is even, $F(h'\,k'\,l') = 2f$.

When $(h' + k' + l')$ is odd, $F(h'\,k'\,l') = 0$. Thus for the bcc lattice, the diffraction is absent for all those planes in which the sum $(h' + k' + l')$ is odd, and is present for the planes in which $(h' + k' + l')$ is even.

Metallic sodium has a bcc structure. The diffraction pattern does not contain lines such as (100), (300), (111), or (221), but lines such as (200), (110), and (222) will be present.

The basis of the fcc structure referred to the cubic cell has identical atoms at $000;\ 0\frac{1}{2}\frac{1}{2};\ \frac{1}{2}0\frac{1}{2};\ \frac{1}{2}\frac{1}{2}0$.

From Eq. (4), the diffraction amplitude becomes
$$F(h'\,k'\,l') = f[1 + e^{\pi i(h'+l')} + e^{\pi i(h'+k')} + e^{\pi i(k'+l')}].$$

Here the expression within the square brackets is the structure factor for *fcc* crystals. The allowed reflections correspond to the cases in which $(h'\,k'\,l')$ are either all even or all odd. No reflections can occur for which the indices are partly even and partly odd.

CRYSTAL BINDING

46.11 Cohesive Energy of Inert Gas Crystals

The Cohesive energy of a crystal is defined as the energy that must be added to the crystal to separate its components into neutral free atoms at rest, at infinite separation, with the same electronic configuration.

The interaction energy of inert gas crystals is due to the following two types of interactions:

(*a*) Van der Waals or London interaction, and

(*b*) Repulsive interaction.

(*a*) Van der Waals-London Interaction

Consider two identical inert gas atoms at a separation R large in comparison with the radii of the atoms (Fig. 46.10). The atoms induce dipole moments in each other. The induced moments cause an attractive interaction between the atoms.

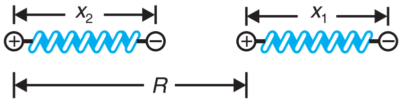
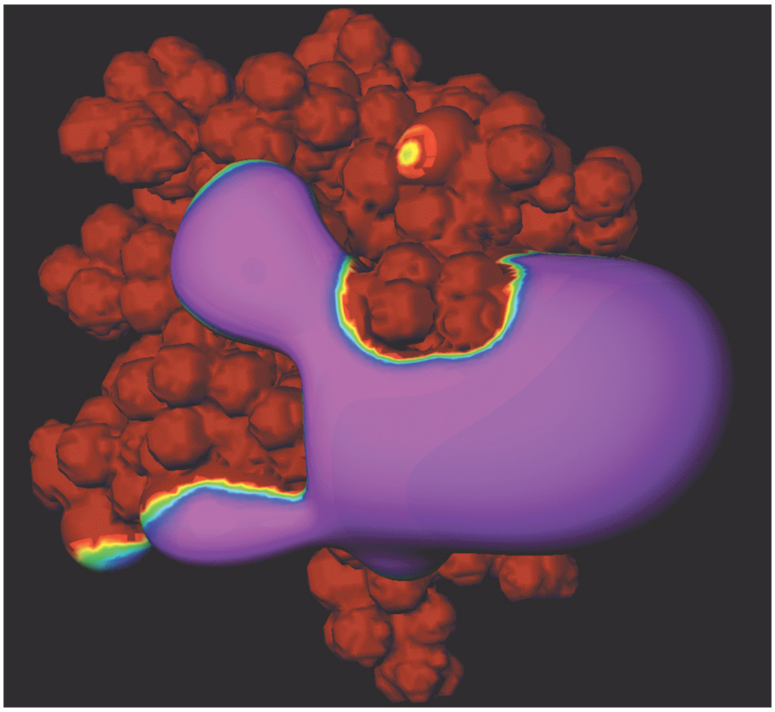


Fig. 46.10

Consider two identical linear harmonic oscillators 1 and 2 separated by R. Each oscillator bears charges $\pm e$ with separations x_1 and x_2. The particles oscillate along the x axis. Let p_1 and p_2 denote the momenta. The force constant is C.



Van der Waals - London Interaction

The hamiltonian of the unperturbed system is

$$H_0 = \frac{1}{2m}p_1^2 + \frac{1}{2}Cx_1^2 + \frac{1}{2m}p_2^2 + \frac{1}{2}Cx_2^2 \qquad \text{...(1)}$$

We assume that each uncoupled oscillator has the frequency ω_0 of the strongest optical absorption line of the atom. Thus $C = m\omega_0^2$ or $\omega_0 = (C/m)^{1/2}$.

Let H_1 be the coulomb interaction energy of the two oscillators. We use CGS units. Then

$$H_1 = \frac{e^2}{R} + \frac{e^2}{R + x_1 - x_2} - \frac{e^2}{R + x_1} - \frac{e^2}{R - x_2} \qquad \text{...(2)}$$

$$H_1 \cong -\frac{2e^2 x_1 x_2}{R^3} \qquad (\because |x_1|, |x_2| \langle\langle R) \quad \text{...(3)}$$

The total hamiltonian $H = H_0 + H_1$ ...(4)

H can be diagonalized by the normal mode transformation

$$x_s \equiv \frac{1}{\sqrt{2}}(x_1 + x_2); \qquad x_a \equiv \frac{1}{\sqrt{2}}(x_1 - x_2). \qquad \text{...(5)}$$

The subscripts s and a denote symmetric and antisymmetric modes of motion. Solving for x_1 and x_2, we get

$$x_1 = \frac{1}{\sqrt{2}}(x_s + x_a); \qquad x_2 = \frac{1}{\sqrt{2}}(x_s - x_a). \qquad \text{...(6)}$$

The momenta p_s, p_a associated with the two modes are,

$$p_1 \equiv \frac{1}{\sqrt{2}}(p_s + p_a); \qquad p_2 \equiv \frac{1}{\sqrt{2}}(p_s - p_a). \qquad \text{...(7)}$$

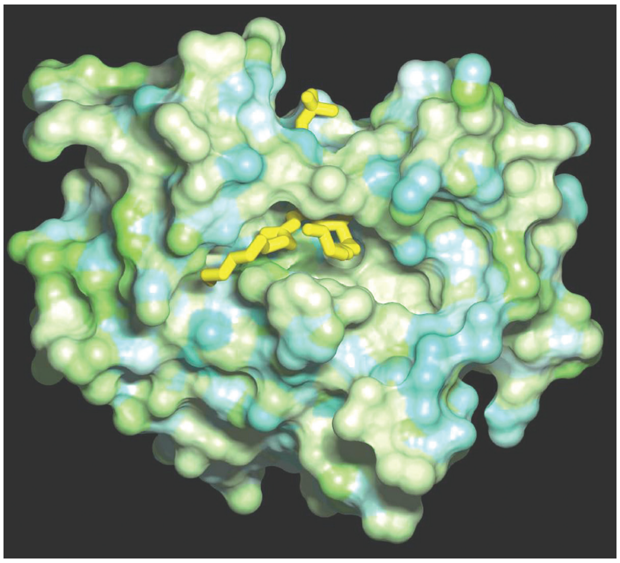
$$H = \left[\frac{1}{2m}P_s^2 + \frac{1}{2}\left(C - \frac{2e^2}{R^3}\right)x_s^2\right] + \left[\frac{1}{2m}P_a^2 + \frac{1}{2}\left(C + \frac{2e^2}{R^3}\right)x_a^2\right] \qquad ...(8)$$

The two frequencies of the coupled oscillators are,

$$\omega = \left[\left(C \pm \frac{2e^2}{R^3}\right)\Big/ m\right]^{1/2} \cong \omega_0\left[1 \pm \frac{1}{2}\left(\frac{2e^2}{CR^3}\right) - \frac{1}{8}\left(\frac{2e^2}{CR^3}\right)^2 + ...\right]. \qquad ...(9)$$

The zero point energy of the system is $\frac{1}{2}\hbar(\omega_s + \omega_a)$. Because of the interaction the sum is lowered from the uncoupled value $2 \cdot \frac{1}{2}\hbar\omega_0$ by

$$\Delta U = \frac{1}{2}\hbar(\Delta\omega_s + \Delta\omega_a) = -\hbar\omega_0 \cdot \frac{1}{8}\left(\frac{2e^2}{CR^3}\right)^2 = -\frac{A}{R^6}. \qquad ...(10)$$



Repulsive Interactions with the Membrane are shown in Blue while Greenish Colours represent the Attractive Interactions.

This attractive interaction varies as the minus sixth power of the separation of the two oscillators.

This is called the van der Waals interaction (or the London interaction or the induced dipole-dipole interaction).

(*b*) Repulsive Interaction

As the two atoms are brought together, their charge distributions gradually overlap. So the electrostatic energy of the system changes. At sufficiently close separations the overlap energy is repulsive, because of the *Pauli exclusion principle.* The electron overlap increases the total energy of the system and gives a repulsive contribution to the interaction. We shall express the repulsive interaction by an empirical repulsive potential of the form B/R^{12}, where **B** is a positive constant.

The total potential energy of two atoms at separation R is

$$U(R) = \frac{B}{R^{12}} - \frac{A}{R^6}$$

or

$$U(R) = 4\varepsilon\left[\left(\frac{\sigma}{R}\right)^{12} - \left(\frac{\sigma}{R}\right)^6\right]. \qquad ...(11)$$

Here, ε and σ are the new parameters, with $4\varepsilon\sigma^6 \equiv A$ and $4\varepsilon\sigma^{12} \equiv B$.

The potential (11) is called *Lennard-Jones Potential* and is plotted in Fig. 46.11.

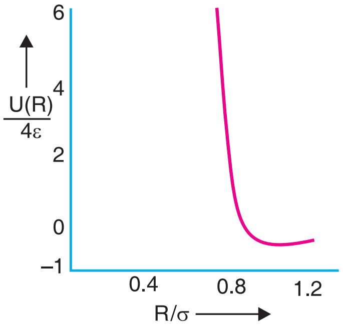


Fig. 46.11

Equilibrium Lattice Constants

We neglect the kinetic energy of the inert gas atoms. The cohesive energy of an inert gas crystal is given by summing the Lennard-Jones potential (11) over all pairs of atoms in the crystal. Suppose that the crystal has N atoms.

The total potential energy is

$$U_{\text{tot}} = \frac{1}{2}N(4\varepsilon)\left[\sum_j{}'\left(\frac{\sigma}{p_{ij}R}\right)^{12} - \sum_j{}'\left(\frac{\sigma}{p_{ij}R}\right)^{6}\right]. \qquad ...(12)$$

Here, $p_{ij}R$ is the distance between reference atom i and any other atom j, expressed in terms of the nearest neighbour distance R.

$\frac{1}{2}N$ has been used instead of N because of the fact that interaction of pair of ij atoms will be only once.

For the fcc structure, $\sum_j{}' p_{ij}^{-12} = 12.13188; \quad \sum_j{}' p_{ij}^{-6} = 14.45392.$ ...(13)

Since there are 12 nearest neighbours of an atom in *fcc* structure, the major contribution to the interaction energy arises from the nearest neighbours.

The equilibrium distance R_0 is obtained from the condition that at $R = R_0$, $dU_{tot}/dR = 0$.

$$\frac{dU_{\text{tot}}}{dR} = -2N\varepsilon\left[(12)(12\cdot13)\frac{\sigma^{12}}{R_0^{13}} - (6)(14\cdot45)\frac{\sigma^{6}}{R_0^{7}}\right]. \qquad ...(14)$$

$$\therefore \qquad R_0/\sigma = 1.09. \qquad ...(15)$$

Cohesive Energy

The cohesive energy of inert gas crystals at absolute zero and at zero pressure is obtained by substituting (13) and (15) in (12).

$$U_{\text{tot}}(R) = 2N\varepsilon\left[(12\cdot13)\left(\frac{\sigma}{R}\right)^{12} - (14\cdot45)\left(\frac{\sigma}{R}\right)^{6}\right]. \qquad ...(16)$$

At $\quad R = R_0, \quad U_{\text{tot}}(R_0) = -(2.15)(4N\varepsilon),$ ...(17)

It is the same for all inert gases.

46.12 Binding Energy of Ionic Crystals

The energy of the crystal is lower than that of the free atoms by an amount equal to the energy required to pull the crystal apart into a set of free atoms. This is called the binding energy (or Cohesive energy) of the Crystal.

The theory developed by Born and Madelung is based on the assumption that the ionic crystals are built up of positive and negative ions each having a spherically symmetric charge distribution. The force between the ions is assumed to be mainly electrostatic. Thus the main contribution to the binding energy arises from electrostatic interaction and is called the *Madelung energy*. The ions tend to acquire such an arrangement in a crystal structure which results in the maximum attractive interaction amongst themselves.

There exist two types of interactions in ionic crystals.

(*i*) The Coulomb's electrostatic interaction energy between two ions with charges ± q is given by

$$\frac{\pm q^2}{4\pi\varepsilon_0 r}$$

Here, r is the distance of separation between them.

(*ii*) There is an interaction due to the overlap of the electronic shells of neighbouring atoms. The repulsive energy is of the form

$$\lambda e^{-r/\rho}.$$

Here, λ represents the strength and ρ the range of repulsive interaction.

This expression for the repulsive energy is called the *central field repulsive potential.*

Let U_{ij} be the interaction energy between ions i and j. Then,

$$U_{ij} = \lambda \exp(-r_{ij}/\rho) \pm \frac{q^2}{4\pi\varepsilon_0 r_{ij}} \qquad ...(1)$$

We define a sum U_i which includes all interactions involving the ion i:

$$U_i = \sum_j{}' U_{ij} \qquad ...(2)$$

Here, the summation includes all ions except $j = i$.

The total lattice energy U_{tot} of a crystal composed of N molecules or $2N$ ions is

$$U_{\text{tot}} = NU_i. \qquad ...(3)$$

Here N, rather than $2N$, occurs because we must count each *pair* of interactions only once. The total lattice energy is defined as the energy required to separate the crystal into individual ions at an infinite distance apart.

We introduce quantities p_{ij} such that $r_{ij} \equiv p_{ij} R$. If ions i and j are the nearest neighbours, then $r_{ij} = R$ or $p_{ij} = 1$.

Here, R is the nearest-neighbour separation in the crystal.

If we include the repulsive interaction only among nearest neighbours, we have

$$U_{ij} = \begin{cases} \lambda \exp(-R/\rho) - \dfrac{q^2}{4\pi\varepsilon_0 R} \text{(nearest neighbours)} \\ \pm \dfrac{1}{p_{ij}} \dfrac{q^2}{4\pi\varepsilon_0 R} \text{(otherwise)} \end{cases} \qquad ...(4)$$

$$\therefore \qquad U_{tot} = NU_i = N\left(z\lambda e^{-R/\rho} - \frac{\alpha q^2}{4\pi\varepsilon_0 R} \right) \qquad ...(5)$$

Here, z is the number of nearest neighbours of any ion and

$$\alpha \equiv \sum_j{}' \frac{(\pm)}{p_{ij}} \equiv \text{Madelung constant} \qquad ...(6)$$

The sum should include the nearest-neighbour contribution, which is just z.

The choice of sign depends upon the type of the reference ion. If the reference ion is negative, the positive sign is used for a positive ion and the negative sign for a negative ion. Thus the value of the Madelung constant depends on the lattice structure. It is basically a correction factor that determines the magnitude of the error introduced by considering only the nearest neighbour interaction.

At the equilibrium separation, R_0, $dU_{tot}/dR = 0$.

$$N\left.\frac{dU_i}{dR}\right|_{R=R_0} = -\frac{Nz\lambda}{\rho}e^{-R_0/\rho} + \frac{N\alpha q^2}{4\pi\varepsilon_0 R_0^2} = 0$$

or
$$R_0^2\, e^{(-R_0/\rho)} = \frac{\rho\alpha q^2}{4\pi\varepsilon_0 z\lambda} \qquad \text{...(7)}$$

This determines the equilibrium separation R_0 if the parameters ρ, λ of the repulsive interaction are known.

The total lattice energy of the crystal of $2N$ ions at their equilibrium separation R_0 may be written, using (5) and (7).

$$U_{tot} = -\frac{N\alpha q^2}{4\pi\varepsilon_0 R_0}\left(1-\frac{\rho}{R_0}\right) \qquad \text{...(8)}$$

The term $\left(\frac{-N\alpha q^2}{4\pi\varepsilon_0 R_0}\right)$ is the *Madelung energy*.

The equilibrium distance R_0 is determined empirically, the range ρ is determined from the knowledge of the bulk modulus of the crystal, and the Madelung constant α is determined theoretically from the geometry of the crystal structure.

Thus the lattice energy of a crystal can be determined.

46.13 Evaluation of the Madelung Constant

The definition of the Madelung constant α is,

$$\alpha = \sum_j{}' \frac{(\pm)}{p_{ij}} \qquad \text{...(1)}$$

An equivalent definition is

$$\frac{\alpha}{R} = \sum_j{}' \frac{(\pm)}{r_j} \qquad \text{...(2)}$$

Here, r_j is the distance of the jth ion from the reference ion and R is the nearest-neighbour distance.

Consider a one-dimensional crystal consisting of alternate positive and negative ions (Fig. 46.12).

Let R be interionic spacing. Take a negative ion as the reference ion. Then

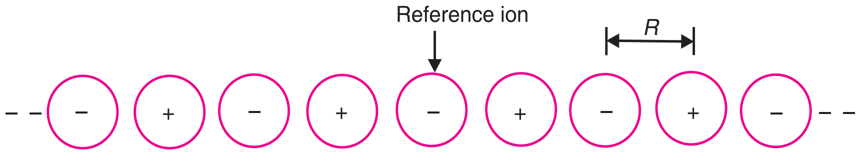


Fig. 46.12

$$\frac{\alpha}{R} = 2\left[\frac{1}{R}-\frac{1}{2R}+\frac{1}{3R}-\frac{1}{4R}+\ldots\right],$$

or
$$\alpha = 2\left[1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\ldots\right] \qquad \text{...(3)}$$

The factor of 2 occurs because there are two equivalent ions, one to the right and one to the left each at distance $r_j = nR$.

Using the series expansion

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots$$

and putting $x = 1$ in it, we get

$$\ln 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots$$

Therefore, Eq. (3) gives

$$\alpha = 2 \ln 2 = 1.38 \qquad \ldots(4)$$

ENERGY BANDS

46.14 Bloch Theorem

Statement. *The solutions of the Schrodinger equation for a periodic potential must be of the form*

$$\psi_{\mathbf{k}}(\mathbf{r}) = u_{\mathbf{k}}(\mathbf{r}) \exp(i\mathbf{k} \cdot \mathbf{r}), \qquad \ldots(1)$$

Here, $u_{\mathbf{k}}(\mathbf{r})$ has the period of the crystal lattice with $u_{\mathbf{k}}(\mathbf{r}) = u_{\mathbf{k}}(\mathbf{r} + \mathbf{T})$.

T is the crystal translation vector.

Explanation. The eigenfunctions of the wave equation for a periodic potential are the product of a plane wave $\exp(i\mathbf{k} \cdot \mathbf{r})$ times a function $u_{\mathbf{k}}(\mathbf{r})$ with the periodicity of the crystal lattice. The function $\psi_{\mathbf{k}}(\mathbf{r})$ in Eq. (1) is called the *Bloch function.*

Proof. We assume that $\psi_{\mathbf{k}}$ is nondegenerate. Consider N identical lattice points on a ring of length Na. The potential energy is periodic in a, with $U(x) = U(x + sa)$.

Here, s is an integer.

The ring has symmetry. We look for solutions of the wave equation such that

$$\psi(x + a) = C\psi(x), \qquad \ldots(2)$$

Here, C is a constant. Then, on going once around the ring,

$$\psi(x + Na) = \psi(x) = C^N \psi(x),$$

because $\psi(x)$ must be single-valued.

So, C is one of the N roots of unity.

$$C = \exp(i2\pi s/N); \qquad s = 0, 1, 2, \ldots, N-1. \qquad \ldots(3)$$

Suppose $u_{\mathbf{k}}(x)$ has the periodicity a, so that $u_{\mathbf{k}}(x) = u_{\mathbf{k}}(x + a)$. Then

$$\psi(x) = u_{\mathbf{k}}(x) \exp(i2\pi sx/Na)$$

with $k = 2\pi s/Na$, satisfies Eq. (2)

Thus the theorem is proved.

46.15 Kronig-Penney Model

This model illustrates the behaviour of electrons in a periodic potential. It assumes a one-dimensional model of periodic potential (Fig. 46.13).

The wave equation can be solved in terms of elementary functions for the square-well array.

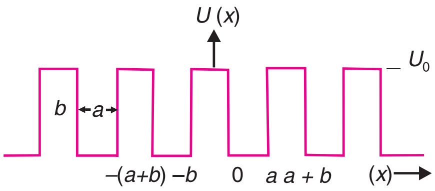


Fig. 46.13

The wave functions associated with this model are calculated by solving the Schrodinger equation

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2}+U(x)\psi = \varepsilon\psi \qquad ...(1)$$

Here, $U(x)$ is the potential energy and ε is the energy eigenvalue.

In the region $0 < x < a$, $U = 0$. Eq. (1) becomes

$$\frac{d^2\psi}{dx^2}+\frac{2m}{\hbar^2}\varepsilon\psi = 0 \qquad ...(2)$$

The eigenfunction is

$$\psi = Ae^{ikx}+Be^{-ikx}, \qquad ...(3)$$

It represents plane waves travelling to the right and to the left, with energy

$$\varepsilon = h^2 K^2/2m \qquad ...(4)$$

In the region $-b < x < 0$, the Schrodinger wave equation is

$$\frac{d^2\psi}{dx^2}+\frac{2m}{\hbar^2}[\varepsilon - U_0]\psi = 0 \qquad ...(5)$$

The eigenfunction is

$$\psi = Ce^{Qx}+De^{-Qx}, \qquad ...(6)$$

and

$$U_0-\varepsilon = h^2 Q^2/2m \qquad ...(7)$$

Since the potential is periodic, the wave functions must be of the form of Bloch function, *i.e.*,

$$\psi(x) = u_k(x)\,e^{ikx} \qquad ...(8)$$

Thus the solution in the region $a < x < a + b$ must be related to the solution (6) in the region $-b < x < 0$ by the Bloch theorem:

$$\psi(a < x < a+b) = \psi(-b < x < 0)\,e^{ik(a+b)}, \qquad ...(9)$$

The constants A, B, C, D are chosen so that ψ and $d\psi/dx$ are continuous at $x = 0$ and $x = a$.

At $x = 0$.

$$A + B = C + D; \qquad ...(10)$$

$$iK(A - B) = Q(C - D). \qquad ...(11)$$

At $x = a$, with the use of (9) for $\psi(a)$ under the barrier in terms of $\psi(-b)$,

$$Ae^{ika}+Be^{-ika} = (Ce^{-Qb}+De^{Qb})\,e^{ik(a+b)}; \qquad ...(12)$$

$$iK(Ae^{ika}-Be^{-ika}) = Q(Ce^{-Qb}-De^{Qb})\,e^{ik(a+b)}. \qquad ...(13)$$

The four equations (10), (11), (12), (13) have a solution only if the determinant of the coefficients of A, B, C, D vanishes. This leads to the following equation:

$$[(Q^2-K^2)/2QK]\;\sinh Qb\sin Ka+\cosh Qb\cos Ka = \cos k(a+b) \qquad ...(14)$$

The result is simplified if we represent the potential by the periodic delta function obtained when we pass to the limit $b = 0$ and $U_0 = \infty$ in such a way that $Q^2 ba/2 = P$, a finite quantity. In this limit $Q >> K$ and $Qb << 1$. Then (14) reduces to

$$(P/Ka)\sin Ka+\cos Ka = \cos ka \qquad ...(15)$$

This is the condition for the solutions of the wave equation to exist.

Figure 46.14 shows a plot of the function $(P/Ka)\sin Ka + \cos Ka$ versus Ka for $P = 3\pi/2$.

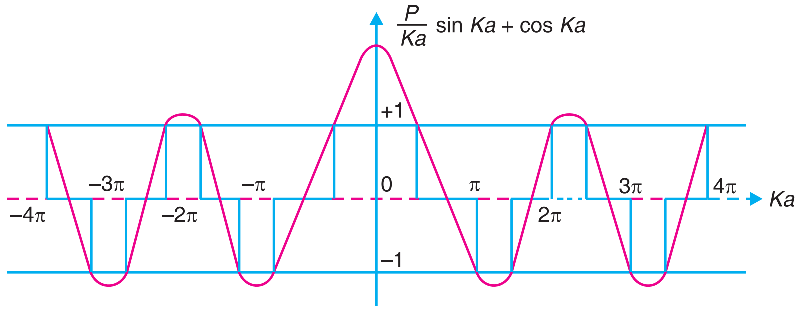


Fig. 46.14

The allowed values of the energy ε are given by those ranges of

$$Ka = \left(\frac{2m\varepsilon}{\hbar^2}\right)^{1/2} a \qquad ...(16)$$

for which the function lies between + 1 and – 1.

Clearly there are regions for Ka where the value of $\left[\frac{P \sin Ka}{Ka} + \cos Ka\right]$ does not lie between – 1 and + 1. For these values of Ka and hence of energy ε, no solutions exist. Such regions of energy are prohibited and are called *forbidden bands.*

Thus, *the energy spectrum of the electron consists of alternate regions of allowed energy* (continuous) and *forbidden energy* (dotted). These regions are usually referred as the allowed and forbidden energy bands.

Fig. 46.15 shows plot of energy vs. wavenumber for the Kronig-Penney potential, with $P = 3\pi / 2$. Notice the energy gaps at $ka = \pi, 2\pi, 3\pi$ ... The figure shows the periodic dependence of ε on k and demonstrates the existence of forbidden energy gaps.

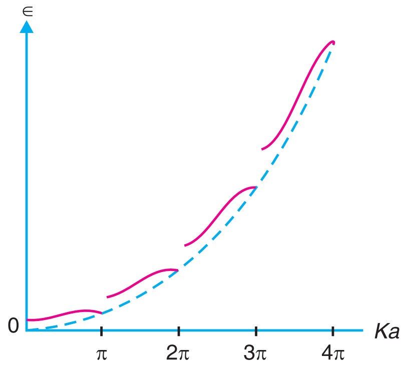


Fig. 46.15

46.16 Effective Mass of Electron

We derive the equation of motion of an electron in an energy band. We look at the motion of a wave packet in an applied electric field E. Suppose that the wave packet is made up of wave functions near a particular wave vector k.

The frequency associated with a wave function of energy ∈ is

$$\omega = \in/\hbar \qquad ...(1)$$

The group velocity is

$$v_g = \frac{d\omega}{dk} = \frac{1}{\hbar}\frac{d\in}{dk} \qquad ...(2)$$

The effects of the crystal on the electron motion are contained in the dispersion relation ∈ (k).

The work $\delta\in$ done on the electron by the electric field E in the interval δt is

$$\delta\in = -eEv_g\,\delta t \qquad ...(3)$$

Using Eq. (2), we observe that

$$\delta\in = (d\in / dk)\,\delta k = \hbar v_g\,\delta k, \qquad ...(4)$$

On comparing Eq.(3) with Eq. (4), we have

$$\delta k = -(eE / \hbar)\delta t, \qquad ...(5)$$

or
$$\hbar\frac{dk}{dt} = -eE \qquad \text{...(6)}$$

We may write Eq.(6) in terms of the external force **F** as
$$\hbar\frac{dk}{dt} = \mathbf{F} \qquad \text{...(7)}$$

In a crystal *hdk / dt* is equal to the external force on the electron.

By differentiating the result of Eq. (2), we get the acceleration $\frac{dv_g}{dt}$ of the wave packet.
$$\frac{dv_g}{dt} = \frac{l}{\hbar}\frac{d^2\in}{dk\,dt} = \frac{l}{\hbar}\left(\frac{d^2\in}{dk^2}\frac{dk}{dt}\right) \qquad \text{...(8)}$$

From Eq. (7) we have
$$dk/dt = F/h \qquad \text{...(8)}$$

Eq.(8) becomes
$$\frac{dv_g}{dt} = \left(\frac{1}{\hbar^2}\frac{d^2\in}{dk^2}\right)F$$

or
$$F = \frac{\hbar^2}{d^2\in/dk^2}\frac{dv_g}{dt} \qquad \text{...(9)}$$

If we identify $\hbar^2/(d^2\in / dk^2)$ as a mass, then Eq. (9) assumes the form of Newton's second law. We define the effective mass m^* by
$$\frac{1}{m^*} = \frac{1}{\hbar^2}\frac{d^2\in}{dk^2} \qquad \text{...(10)}$$

An electron in a periodic potential is accelerated relative to the lattice in an applied electric or magnetic field as if the mass of the electron were equal to an effective mass m, defined by Eq.(10).*

It is easy to generalize Eq. (10) to take account of an anisotropic energy surface, as for electrons in Si or Ge. We introduce the components of the reciprocal *effective mass tensor*:
$$\left(\frac{1}{m^*}\right)_{\mu\nu} = \frac{1}{\hbar^2}\frac{d^2\in_k}{dk_\mu\,dk_\nu}\,; \quad \frac{dv_\mu}{dt} = \left(\frac{1}{m^*}\right)_{\mu\nu}F_\nu \qquad \text{...(1)}$$

Here, μ, v are Cartesian coordinates.

Effective Mass Tensor. In a crystal, for an electron the ratio of force and acceleration is generally a tensor of second order. It is called the *effective mass tensor*. Eq. (11) is the formula of effective mass tensor for nearly free electron model. The formula would be valid for real potentials for an electron in a crystal. In simplest situations this tensor is transformed into a scalar.

Nature of Effective Mass. For the electron in a crystal the value of m^* depends on its energy (Fig. 46.16). Fig. 46.16 shows the variation of energy ∈ with wave vector k. The effective mass m^* is determined by the curvature $\frac{d^2\in}{dk^2}$.

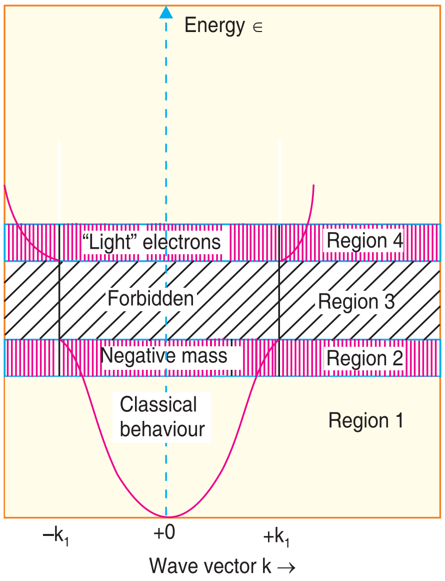


Fig. 46.16

(*i*) In region 1, the parabolic relation is good, and behaviour is therefore classical.

(*ii*) In region 2, the curvature is negative. Therefore, the effective mass is *negative*.

(*iii*) In region 4, the curvature is positive but larger than that of region 1. Therefore, the effective mass is *positive*. The effective mass is smaller, according to Eq. (10).

Physical Interpretation of the Effective Mass

A single electron in an energy band may have positive or negative effective mass m^*.

(*i*) The states of positive effective mass occur near the bottom of a band because positive effective mass means that the band has upward curvature ($d^2\in / dk^2$ is positive).

(*ii*) States of negative effective mass occur near the top of the band. A negative effective mass means that on going from state k to state $k + \Delta k$, the momentum transfer to the lattice from the electron is larger than the momentum transfer from the applied force to the electron.

(*iii*) If the energy in a band depends only slightly on k, then the effective mass will be very large. That is, $m^*/m > 1$ when $d^2\in / dk^2$ is very small.

Explanation of negative masses which occur near a Brillouin zone boundary :

(*i*) In Fig. 46.17 (*a*), the energy of the electron beam incident on a thin crystal is slightly too low to satisfy the condition for Bragg reflection and the beam is transmitted through the crystal.

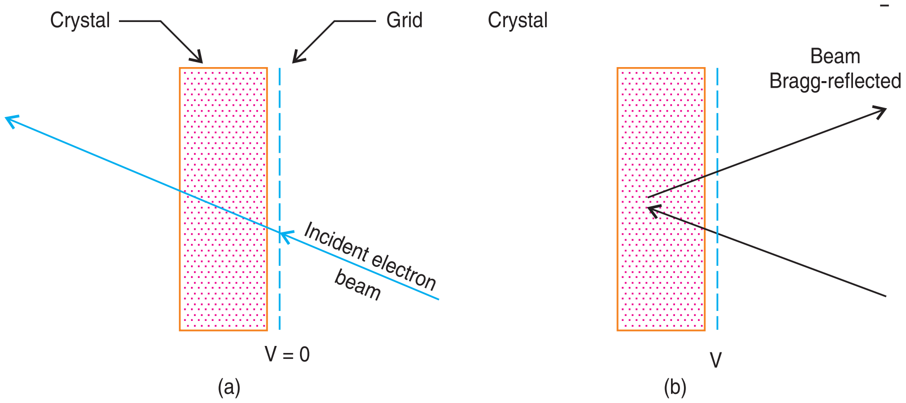


Fig. 46.17

(*ii*) In Fig. 46.17 (*b*), a small accelerating potential applied on the grid leads to the required condition for Bragg reflection. Then the electron beam will be reflected by appropriate set of crystal planes. Thus, a small change in energy of electron beam would induce a large change in the momentum associated with the small effective mass. If the grid potential accelerates an electron beam of small energy, the effective mass is negative (Because the acceleration occurs in a direction opposite to change in momentum).

INTERATOMIC POTENTIAL

46.17 Interatomic Potential

The potential between a pair of atoms is called the *interatomic potential*.

From the very existence of solids we draw the following two general conclusions:

(*i*) There must act attractive forces between the atoms or molecules in a solid which keep them together.

(*ii*) There must be repulsive forces acting between the atoms as well, since large external pressures are required to compress a solid to appreciable extent.

Consider a single pair of atoms A and B which form a stable chemical compound.

Assume that the potential energy of atom B due to the presence of atom A is given by

$$U(r) = -\frac{\alpha}{r^n} + \frac{\beta}{r^m} \qquad ...(1)$$

Here r is the distance between the nuclei of the two atoms. α, β, m and n are constants characteristic for the molecule AB.

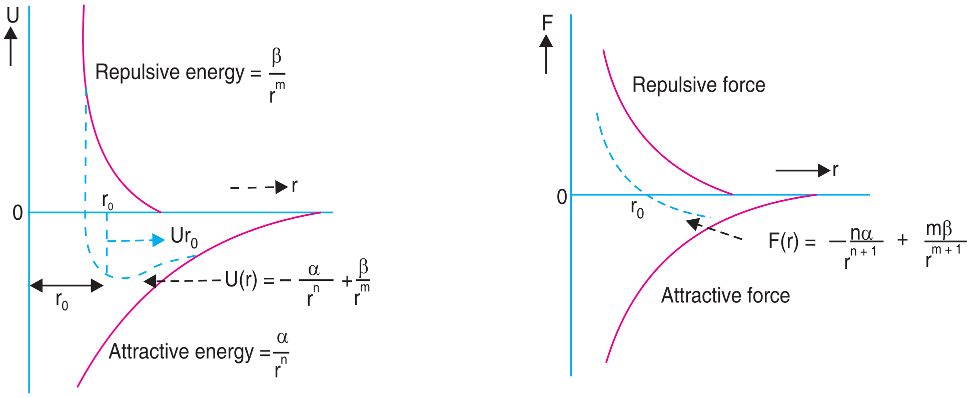
The first term, $\frac{\alpha}{r^n}$, which is negative, corresponds to the energy associated with the forces of attraction.

The second term, $\frac{\beta}{r^n}$, which is positive, corresponds to the forces of repulsion.

The force between the two atoms A and B as function of r is given by

$$F(r) = -\frac{dU}{dr} = -\frac{n\alpha}{r^{n+1}} + \frac{m\beta}{r^{m+1}} \qquad ...(2)$$

The energy (U) and force (F) between two atoms A and B as function of their separation r are plotted in Fig. 46.18. Dotted curves are the sums of the attractive and repulsive curves.



(a) (b)

Fig. 46.18

The attraction between two atoms brings them closer together until the individual electron clouds begin to overlap and a strong repulsive force arises to comply with Pauli's exclusion principle. When the attractive force and the repulsive force between any two atoms are equal, the two atoms should be in a stable situation with a minimum potential energy corresponding to the equilibrium interatomic distance, r_0.

At the equilibrium separation $r = r_0$, the potential energy must be a minimum and the first derivative of Eq. (1) must vanish.

$$\left.\frac{dU}{dr}\right|_{r=r_0} = 0 = \frac{n\alpha}{r^{n+1}} - \frac{m\beta}{r^{m+1}} \qquad ...(3)$$

$$\therefore \qquad \beta = \alpha.\frac{n}{m}(r_0)^{m-n} \qquad ...(4)$$

and $$(r_0)^{m-n} = \left(\frac{\beta}{\alpha}\right)\left(\frac{m}{n}\right) \qquad ...(5)$$

According to Eq. (2), this condition is equivalent to the requirement that the attractive and repulsive forces balance, *i.e.*, $F(r_0) = 0$.

From Eqs. (1) and (4), we obtain for the energy in the equilibrium state

$$U(r_0) = -\frac{\alpha}{r_0^n} + \frac{\beta}{r_0^m}$$

$$\therefore \qquad U(r_0) = -\frac{\alpha}{r_0^n}\left(1 - \frac{n}{m}\right) \qquad ...(6)$$

The energy corresponding to $r = r_0$ position, denoted by $U(r_0)$ is called the binding energy or the energy of the cohesion of the molecule. This is the energy required to remove the two atoms of the molecule (AB) into an infinite separation.

Dissociation Energy. The positive quantity $D = -U(r_0)$ is the *dissociation energy* of the molecule, *i.e.,* the energy required to separate the two atoms. Dissociation may occur, for example, at high temperatures or as a result of other processes in which the molecule can absorb sufficient energy. The dissociation energies are of the order of 1eV to a few eV.

A minimum in the energy curve is possible only if $m > n$.

Thus *the formation of a chemical bond requires that the repulsive forces be of shorter range than the attractive forces.*

This may be shown by employing the condition for energy minimum,

$$\left.\frac{d^2U}{dr^2}\right|_{r=r_0} > 0 \qquad ...(7)$$

This condition leads to

$$\frac{-n(n+1)\alpha}{r_0^{n+2}} + \frac{m(m+1)\beta}{r_0^{m+2}} > 0 \qquad ...(8)$$

Substituting the value of r_0 from Eq. (5) in Eq. (8), we obtain

$$m > n. \qquad ...(9)$$

EXAMPLE 1. *If the potential energy function is expressed as*

$$U(r) = -\frac{\alpha}{r^6} + \frac{\beta}{r^{12}}$$

show that

(*a*) *The intermolecular distance* r_0 *for which the potential energy is minimum is given by* $\left(\frac{2\beta}{\alpha}\right)^{1/6}$.

(*b*) *The minimum potential energy is given by,* $U_{min} = -\frac{\alpha^2}{4\beta}$.

SOL. (*a*) The condition for the minimum in the potential energy is

$$\left(\frac{dU}{dr}\right)_{at\, r=r_0} = 0$$

$$\therefore \qquad \left(\frac{dU}{dr}\right)_{r=r_0} = 0 = \left[\frac{6\alpha}{r^7} - \frac{12\beta}{r^{13}}\right]_{r=r_0}$$

$$\text{or} \qquad \frac{6\alpha}{r_0^7} = \frac{12\beta}{r_0^{13}} \text{ at } r = r_0$$

$$\text{or} \qquad r_0 = \left(\frac{2\beta}{\alpha}\right)^{1/6} \qquad ...(1)$$

(*b*) The minimum potential energy is

$$U_{min} = \left[-\frac{\alpha}{r_0^6} + \frac{\beta}{r_0^{12}}\right]$$

$$= -\frac{\alpha}{r_0^6}\left[1 - \frac{\beta}{\alpha}\frac{1}{r_0^6}\right]$$

$$= -\frac{\alpha}{\left(\frac{2\beta}{\alpha}\right)}\left[1 - \frac{\beta}{\alpha}.\frac{\alpha}{2\beta}\right] \qquad \text{using Eq. (1)}$$

$$U_{min} = -\frac{\alpha^2}{2\beta}\left[1 - \frac{1}{2}\right] = -\frac{\alpha^2}{4\beta}$$

46.18 Second Order Expansion of Potential Energy

The total potential energy of two atoms at separation r is

$$U = -\frac{\alpha}{r^n} + \frac{\beta}{r^m} \qquad ...(1)$$

Here, α, β, n and m are constants.

Let r_0 be the equilibrium interatomic distance.

Then, $(r - r_0)$ is the displacement from equilibrium position.

It is convenient to write

$$r = r_0 + (r - r_0) \qquad ...(2)$$

$$\therefore \qquad U = -\frac{\alpha}{[r_0 + (r - r_0)]^n} + \frac{\beta}{[r_0 + (r - r_0)]^m} \qquad ...(3)$$

$$= -\left(\frac{\alpha}{r_0^n}\right)\frac{1}{\left[1 + \left(\frac{r - r_0}{r_0}\right)\right]^n} + \left(\frac{\beta}{r_0^m}\right)\frac{1}{\left[1 + \left(\frac{r - r_0}{r_0}\right)\right]^m}$$

$$U = -\left(\frac{\alpha}{r_0^n}\right)\left[1 + \left(\frac{r - r_0}{r_0}\right)\right]^{-n} + \frac{\beta}{r_0^m}\left[1 + \left(\frac{r - r_0}{r_0}\right)\right]^{-m} \qquad ...(4)$$

In the second-order expansion, we can write

$$U = -\frac{\alpha}{r_0^n}\left[1 - n\left(\frac{r - r_0}{r_0}\right) - \frac{n(-n-1)}{2}\left(\frac{r - r_0}{r_0}\right)^2\right]$$

$$+\frac{\beta}{r_0^m}\left[1-m\left(\frac{r-r_0}{r_0}\right)-\frac{m(-m-1)}{2}\left(\frac{r-r_0}{r_0}\right)^2\right]$$

$$U = \left[-\frac{\alpha}{r_0^n}+\frac{\beta}{r_0^m}\right]+\left[\frac{\alpha n}{r_0^n}-\frac{\beta m}{r_0^m}\right]\left(\frac{r-r_0}{r_0}\right)$$

$$+\left[-\frac{\alpha}{r_0^n}\frac{n(n+1)}{2}+\frac{\beta}{r_0^m}\frac{m(m+1)}{2}\right]\left(\frac{r-r_0}{r_0}\right)^2 \qquad ...(5)$$

Let $-\frac{\alpha}{r_0^n}+\frac{\beta}{r_0^m} = U_0$ = equilibrium value of interatomic potential energy, $\frac{1}{r_0}\left[\frac{\alpha n}{r_0^n}-\frac{\beta m}{r_0^m}\right] = A$,

and $\frac{1}{2r_0^2}\left[-\frac{\alpha n(n+1)}{r_0^n}+\frac{\beta m(m+1)}{r_0^m}\right] = B$.

Eq. (5) becomes,

$$U = U_0 + A(r-r_0) + B(r-r_0)^2 \qquad ...(6)$$

The force between the two atoms is given by

$$F = -\frac{dU}{dr} = -[A + 2B(r-r_0)] \qquad ...(7)$$

At equilibrium, $F = 0$. $\therefore$ $A = 0$.

$$F = -2B(r-r_0)$$

Let $2B = k$ = constant.

$$\therefore \qquad F = -k(r-r_0) \qquad ...(8)$$

Eq. (8) means that if we pull the atoms apart and let them go, they will oscillate about the equilibrium position. That is, a restoring force would come into play.

A motion governed by the force of Eq. (8) is called the *harmonic motion*.

46.19 Second Order Expansion of Lennard-Jones Potential about the Minimum

Lennard-Jones Potential. This interatomic potential is widely used in theoretical consideration of the physical properties of the materials. It is expressed as

$$U = -\frac{C}{r^6}+\frac{D}{r^8} \qquad ...(1)$$

Here, r is distance between the atoms.

C and D are constants.

Let r_0 be the equilibrium interatomic distance.

Then, $(r-r_0)$ is the displacement from equilibrium position. It is convenient to write

$$r = r_0 + (r-r_0) \qquad ...(2)$$

$$\therefore \qquad U = -\frac{C}{[r_0+(r-r_0)]^6}+\frac{D}{[r_0+(r-r_0)]^8} \qquad ...(3)$$

$$U = -\frac{C}{r_0^6}\left[1+\left(\frac{r-r_0}{r_0}\right)\right]^{-6}+\frac{D}{r_0^8}\left[1+\left(\frac{r-r_0}{r_0}\right)\right]^{-8}$$

In the second order approximation,

$$U = -\frac{C}{r_0^6}\left[1-6\left(\frac{r-r_0}{r_0}\right)+\frac{42}{2}\left(\frac{r-r_0}{r_0}\right)^2\right]$$

$$+\frac{D}{r_0^8}\left[1-8\left(\frac{r-r_0}{r_0}\right)+\frac{72}{2}\left(\frac{r-r_0}{r_0}\right)^2\right]$$

$$U = \left[-\frac{C}{r_0^6}+\frac{D}{r_0^8}\right]+\frac{1}{r_0}\left[\frac{6C}{r_0^6}-\frac{8D}{r_0^8}\right](r-r_0)+\frac{1}{2r_0^2}\left[-\frac{42C}{r_0^6}+\frac{72D}{r_0^8}\right](r-r_0)^2$$

Let $-\frac{C}{r_0^6}+\frac{D}{r_0^8}=U_0$ = equilibrium value of interatomic potential energy,

$$\frac{1}{r_0}\left[\frac{6C}{r_0^6}-\frac{8D}{r_0^8}\right] = A$$

and

$$\frac{1}{2r_0^2}\left[-\frac{42C}{r_0^6}+\frac{72D}{r_0^8}\right] = B$$

$$\therefore \qquad U = U_0 + A(r-r_0) + B(r-r_0)^2 \qquad ...(4)$$

The force between the two atoms is given by

$$F = -\frac{dU}{dr} = -[A + 2B(r-r_0)] \qquad ...(5)$$

At equilibrium, $F = 0. \quad \therefore \quad A = 0.$

$$F = -2B(r-r_0)$$

Let $2B=\frac{6}{r_0^8}\left[-7C+\frac{12D}{r_0^2}\right]=k$ = spring constant.

$$\therefore \qquad F = -k(r-r_0) \qquad ...(6)$$

Eq. (6) means that if we pull the atoms apart and let them go, they will oscillate about the equilibrium position. That is, a restoring force would come into play. A motion governed by the force of Eq. (6) is called the *harmonic motion*.

Let μ be the reduced mass of the molecule, say, *NaCl. The characteristic circular frequency of the molecule is*

$$\omega = \left(\frac{k}{\mu}\right)^{1/2} \qquad ...(7)$$

46.20 Simple Harmonic Oscillator

Lattice Vibrations. Lattice dynamics concerns itself with the vibrations of atoms about their equilibrium sites in a solid. These vibrations occur at any temperature, even at absolute zero.

A convenient way to visualise lattice vibrations is to consider a large number of spheres, each connected to its neighbours by springs. These springs provide restoring forces on the spheres when they are displaced from their equilibrium positions. Our discussion will be entirely based on the "harmonic model". We remain in the linear region (or region of elastic deformation) such that the restoring force on each atom is approximately proportional to the displacement (*Hooke's law*). Springs connecting only nearest neighbours characterise short range forces. If each of the particles represented by the sphere is imagined to be vibrating about its equilibrium position, a fairly accurate picture of lattice vibrations emerges.

PHOTOGATE
3523

Simple Harmonic Oscillator.

The equilibrium positions of the atoms are determined by a balance between the attractive and repulsive forces. At equilibrium the potential energy of the solid must be a minimum.

The potential energy of an atom, E_a, in a solid containing N atoms, is sketched as a function of the lattice spacing of the solid in Fig. 46.19.

(*i*) For large lattice spacing the potential energy is zero, since the atoms then interact negligibly with each other.

(*ii*) As the lattice spacing decreases, the potential energy decreases. Finally, at some critical spacing the potential energy passes through a minimum.

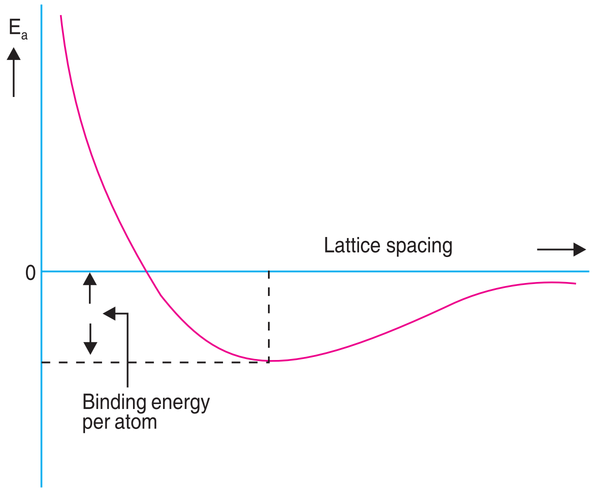
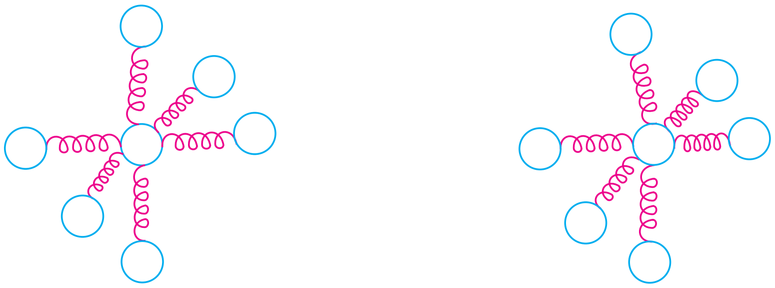


Fig. 46.19

(*iii*) As the lattice spacing further decreases, the potential energy increases rapidly.

The important region is that near the minimum, where no net force is exerted on an atom by its neighbours. In real crystals, there are forces in all three directions, and the motion of each atom is three-dimensional. Here, we study a *one-dimensional* situation.



(a) (b)

Fig. 46.20

All the atoms of the solid except one are fixed in position. Fig. 46.18 (a) shows the *atom in its equilibrium position*.

The atom in question moves in the x direction against its neighbours. Fig. 46.20 (b) shows the *atoms displaced a distance u in the x direction*. Then, unbalanced forces (represented by the extension and compression of springs) develop between the neighbouring atoms along the x axis.

The springs in the y and z directions do not change their length. Since the distance to neighbours in the y and z directions is assumed to be unchanged, the unbalanced force in these directions is zero.

Frequency of Vibration

Fig. 46.19 represents the energy change when the distances to all the neighbours of an atom vary. If only one varies, the energy $E_a(x)$ must be divided by z, where z is the number of nearest neighbours. (We assume that the potential energy for atoms farther away than the nearest neighbour can be neglected).

The change in energy, ΔE, which accompanies a change in atom position from x_0 to x is

$$\Delta E = \frac{2}{z}\{E_a(x_0 - u) + E_a(x_0 + u) - 2\,E_a(x_0)\} \qquad ...(1)$$

Here, $u = x - x_0$

The first term is the energy of the bond with the right neighbour and the second is the energy of the bond with the left. The factor 2 appears because the movement of an atom against its neighbour causes both the energy of the displaced atom and that of its neighbour to change by an equal amount.

The function $E_a(x)$ can be expanded in a Taylor series about its minimum.

The total energy change of the atom and its neighbours for small displacements is

$$\Delta E = \frac{4}{2z}\left(\frac{\partial^2 E_a}{\partial x^2}\right)_{x_0}(x - x_0)^2 = \frac{k}{2}u^2 \qquad ...(2)$$

Here, $u = x - x_0$ and $k = \frac{4}{z}\left(\frac{\partial^2 E_a}{\partial x^2}\right)_{x_0}$.

k represents the *force constant*.

An atom bound to a definite site by the potential-energy law of Eq. (2) is a simple harmonic oscillator. Hence, the vibrating atoms execute harmonic oscillations about their equilibrium positions.

The force F on each atom as a function of its displacement u is given by

$$F = -\frac{d(\Delta E)}{du} = -ku \qquad ...(3)$$

Let m be the mass of the atom.

The differential equation describing the motion is

$$m\frac{d^2u}{dt^2} = -ku \quad ...(4)$$

A solution of Eq. (4) is

$$u = A_0 \cos\left(\frac{k}{m}\right)^{1/2} t \quad ...(5)$$

The angular frequency ω of the atom is given by

$$\omega = \left(\frac{k}{m}\right)^{1/2}$$

The frequency of vibration (called the *Einstein frequency*) is given by

$$\nu = \frac{1}{2\pi}\left(\frac{k}{m}\right)^{1/2} \quad ...(6)$$

The force constant k of the atom is approximately 25 Nm^{-1}. For atoms in the middle of the periodic table, *Cu*, *Fe*, *Ge*, etc., the atomic masses are about 10^{-25} *kg*. The frequency of vibration, ν of these atoms is $\approx 3 \times 10^{12}$ Hz. It can be shown that the amplitude of oscillation (at room temperature) is a few hundredths of an angstrom unit, *i.e.*, a few per cent of the interatomic spacing. Each separate mechanical oscillation at a given frequency is called a *phonon*. Each *phonon* has an energy (*i.e.,* amplitude of oscillation) as well as a frequency.

Review question. Define a simple harmonic oscillator. By expanding the potential energy of an atomic oscillator, derive a formula for the Einstein frequency.

EXERCISE

1. Explain the concept of reciprocal lattice. Discuss its properties. What is its importance?
2. Derive the Bragg conditions in terms of the reciprocal lattice vectors. (*Raj*., 1974)
3. Explain the concept of Brillouin zones. Give the construction of the first Brillouin zone of an oblique lattice in two dimensions.
4. Show that the reciprocal lattice of a simple cubic lattice is another simple cubic lattice. Hence deduce the first Brillouin zone of the simple cubic crystal lattice.
5. Show that the reciprocal lattice for a body centred cubic crystal is face centred cubic. (*Raj*., 1974; *Jodhpur*, 1968)
6. What do you mean by reciprocal lattice? Calculate the reciprocal lattice for a direct *fcc* lattice. What is meant by first Brillouin Zone? (*Delhi*, 1983)
7. Discuss the origin of the geometrical structure factor and the atomic scattering factor in x-ray diffraction and derive suitable mathematical expressions for them. What is the importance of the geometrical structure factor in the analysis of the crystal structures? (*Raj*., 1977)
8. Explain the origin of van der Waals forces in molecular crystal. Show that for van der Waals forces the interaction energy varies as $1 / R^6$, where R is the separation of two interacting atoms. (*Raj*., 1984)
9. Define Cohesive energy and determine its value for crystals of inert gases.
10. Obtain the binding energy of an ionic crystal and derive the expression for the Madelung constant. Calculate the value of this constant for a linear ionic lattice. (*Raj*., 1970)
11. Obtain an expression for the total cohesive energy of an ionic crystal in terms of Madelung constant and other parameters.
12. State and prove Bloch theorem.

13. Discuss the Kronig-Penney model for a linear lattice. How does it lead to formation of energy bands in solid ? (*Raj*., 1983)

14. Give Kronig -Penney model of electron in a periodic potential. What are its consequences?

(*Rohilkhand,* 1993, 92, 91; *Meerut,* 1993)

15. Define "Reciprocal lattice".

Ans. Consider a set of points **R** constituting a Bravais lattice, and a plane wave, $e^{i\mathbf{k}.r}$. For general **k,** such a plane wave will not have the periodicity of the Bravais lattice, but for certain special choices of wave vector it will. The set of all wave vectors **K** that yield plane waves with the periodicity of a given Bravais lattice is known as its reciprocal lattice. Analytically, **K** belongs to the reciprocal lattice of a Bravais lattice of points ***R,*** provided that the relation

$$e^{i\mathrm{K}\cdot(r+R)} = e^{i\mathrm{K}\cdot\mathrm{r}} \qquad ...(1)$$

holds for any **r**, and for all **R** in the Bravais lattice. Factoring out $e^{iK.r}$, we can characterize the reciprocal lattice as the set of wave vectors **K** satisfying

$$e^{i\mathrm{K}\cdot\mathrm{R}} = 1 \qquad ...(2)$$

for all **R** in the Bravais lattice.

Note that a reciprocal lattice is defined with reference to a particular Bravais lattice. The Bravais lattice that determines a given reciprocal lattice is often referred to as the *direct lattice*, when viewed in relation to its reciprocal. Note also that although one could define a set of vectors **K** satisfying (2) for an arbitrary set of vectors **R**, such a set of **K** is called a reciprocal lattice only if the set of vectors **R** is a Bravais lattice.

CHAPTER

47

MAGNETIC PROPERTIES OF MATERIALS

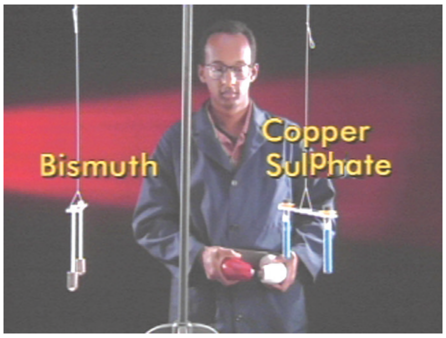
AT A GLANCE			
47.1	Introduction	**47.2**	Langevin's Theory of Diamagnetism
47.3	Langevin's Theory of Paramagnetism	**47.4**	Ferromagnetism
47.5	Soft and Hard Magnetic Materials		

47.1 Introduction

According to their magnetic properties, magnetic materials are divided into three principal groups: *diamagnetic*, *paramagnetic* and *ferromagnetic* materials. The vector quantity **M**, called the *intensity of magnetisation*, is used to characterise the magnetisation of substances. It is the vector sum of the magnetic moments of the atoms (or molecules) contained in unit volume. Thus

$$\mathbf{M} = \lim_{V \to 0} \left(\frac{1}{V} \sum_{i=1}^{N} \mathbf{P}_{mi} \right)$$

where N = Number of particles in volume V of the magnetic material, P_{mi} = magnetic moment of the ith atom (or molecule). For a large class of magnetic materials it is found that *intensity of magnetisation* (**M**) is directly proportional to *magnetic field intensity* (**H**),



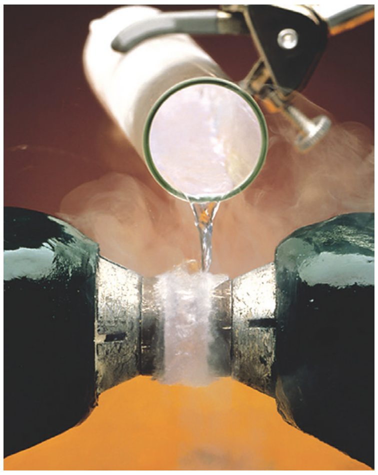
Bismuth is Weakly Repelled and Copper Sulphate Weakly Attracted by a Strong Magnet.

$$i.e., \mathbf{M} \propto \mathbf{H} \quad \text{or} \quad \mathbf{M} = \chi\, \mathbf{H}$$

where χ is known as the *magnetic susceptibility* of the substance.

Diamagnetism. In 1846 Michael Faraday discovered that a specimen of bismuth brought near to the pole of a strong magnet is *repelled.* He called such substances *diamagnetic.* Antimony, bismuth, mercury, gold and copper are some examples of diamagnetic substances. For diamagnetic substances, $\chi < 0$ and is independent of temperature. Unlike *paramagnetic* materials, whose atoms or molecules have a net magnetic moment, the atoms or molecules of a diamagnetic material have zero magnetic moment in the absence of an external applied field. With the application of an external magnetic field a net dipole moment, opposing the field, is induced in the atoms or molecules. In paramagnetic materials there is also a diamagnetic effect but it is much weaker than that of the paramagnetism.

Paramagnetism. *Paramagnetic substances* are attracted towards the region of stronger magnetic field. Aluminium, oxygen and platinum are some examples of paramagnetic substances. Their magnetic susceptibility $\chi > 0$ and depends on temperature according to the Curie law: $\chi = C/T$ where C is called Curie's constant.



Paramagnetism of Oxygen

Paramagnetism occurs in those substances where the individual atoms, ions, or molecules possess a permanent magnetic dipole moment. In the absence of an external magnetic field, the atomic dipoles point in random directions. So there is no resultant magnetisation of the substance as a whole in any direction. This random orientation is the result of thermal agitation within the substance. When an external field is applied, the atomic dipoles tend to orient themselves parallel to the field, since this is a state of lower energy than the antiparallel position. This gives a net magnetisation parallel to the field, and a positive contribution to the susceptibility. At low temperatures, the thermal agitation, which tends to give a random orientation to the atomic dipoles, is less. Hence a bigger proportion of the dipoles are able to align themselves parallel to the field, and the magnetisation is greater for a given field. For ordinary fields and temperatures, $\chi \propto 1/T$. For large fields at low temperatures the magnetisation produced is no longer proportional to the applied field, and tends to a constant value. This saturation effect is produced when all the atomic dipoles are aligned parallel to the field, so that the magnetisation reaches a limiting maximum value.

Ferromagnetism. *Ferromagnetic materials* are very strongly magnetic. For a given field, magnetisation in ferromagnetic substances is about 10^6 times more than that for dia or paramagnetic substances. Iron, nickel and cobalt are some examples of ferromagnetic substances. As temperature increases, the value of χ decreases. Above a certain temperature, known as Curie temperature, ferromagnetics become paramagnetics.

47.2 Langevin's Theory of Diamagnetism

Consider an electron (mass = m, charge = e) rotating about the nucleus (charge = Ze) in a circular orbit of radius r. Let ω_o be the angular velocity of the electron. Then

$$F_o = m\omega_0^2\, r = Ze^2/(4\pi \in_0 r^2)$$

or

$$\omega_0 = \sqrt{\frac{Ze^2}{4\pi \in_0 mr^3}} \qquad \text{...(1)}$$

The magnetic moment of the electron is

$$\vec{m} = \text{current} \times \text{area} = \frac{e\omega_0}{2\pi} \times \pi r^2 = \frac{e}{2}\,\omega_0\, r^2 \qquad ...(2)$$

Let a magnetic field of induction ***B*** be now applied. ***B*** is normal to and into the page (Fig. 47.1).

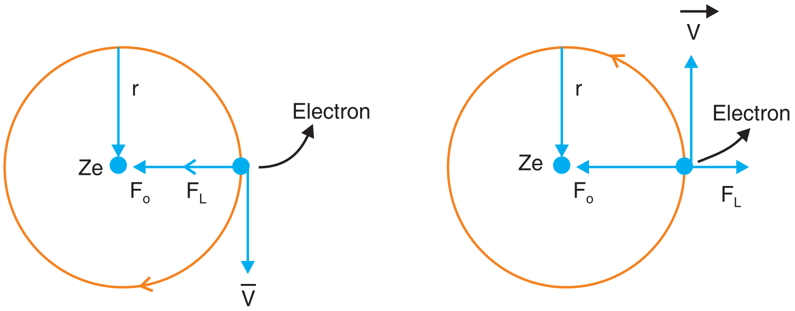


Fig. 47.1

An additional force F_L called Lorentz force, acts on the electron.

$$\mathbf{F}_L = -e\,(\mathbf{v} \times \mathbf{B}) = -\,eBr\omega$$

The condition of stable motion is now given by

$$mr\omega^2 = \frac{Ze^2}{4\pi \in_0 r^2} - eBr\omega \qquad ...(3)$$

or

$$\omega^2 + \frac{eB}{m}\omega - \frac{Ze^2}{4\pi \in_0 mr^3} = 0$$

Solving the quadratic equation in ω,

$$\omega = \frac{\dfrac{-eB}{m} \pm \sqrt{\left(\dfrac{eB}{m}\right)^2 + 4\left(\dfrac{Ze^2}{4\pi \in_0 mr^3}\right)}}{2} = \pm\sqrt{\omega_0^2 + \left(\frac{eB}{2m}\right)^2} - \frac{eB}{2m}.$$

or

$$\omega = \pm\omega_0 - \frac{eB}{2m} \qquad \left(\because \frac{eB}{2m} << \omega_0\right) \qquad ...(4)$$

Thus the angular frequency is now different from ω_o. The result of establishing a field of flux density ***B*** is to set up a precessional motion of the electronic orbits with angular velocity $-(e/2m)$ ***B***. This is called Larmor theorem. Then

$$\left.\begin{array}{l}\text{change in frequency of}\\ \text{revolution of the electron}\end{array}\right\} = \delta n = -\frac{eB}{4\pi m}$$

The corresponding change in the magnetic moment of the electron is

$$\Delta m = \text{current} \times \text{area} = \left\{e \times \left(\frac{-eB}{4\pi m}\right)\right\} \times \pi r^2 = -\frac{Be^2 r^2}{4m} \qquad ...(5)$$

On summing over all electrons in the atom, the induced moment per atom becomes

$$\Delta m_{atom} = -\frac{Be^2 \Sigma r^2}{4m}$$

Let N be the number of atoms per unit volume. Then the magnetisation M is given by

$$M = -\frac{NBe^2 \Sigma r^2}{4m} \qquad ...(6)$$

All the electron orbits are not oriented normal to the magnetic field. Hence r^2 in Eq. (6) should be replaced by the average of the square of the projection of orbit radii for various electrons in a plane perpendicular to B. Hence we should replace r^2 in Eq. (6) by $\frac{2}{3}r^2$. Therefore,

$$M = -\frac{NBe^2\Sigma r^2}{6m}$$

Volume susceptibility of the material

$$\chi = \frac{M}{H} = -\frac{NBe^2\Sigma r^2}{6mH} = -\frac{\mu_0 Ne^2\Sigma r^2}{6m} \qquad (\because B = \mu_0 H)$$

$$\therefore \quad \chi = -\frac{\mu_0 Ne^2\Sigma r^2}{6m} = -\frac{\mu_0 e^2}{6m} NZ <r^2> \qquad ...(7)$$

Eq. (7) shows that χ is independent of the field strength and temperature. This is in accord with Curie's experimental results.

47.3 Langevin's Theory of Paramagnetism

He assumes that each atom has a permanent magnetic moment m. The only force acting on the atom is that due to the external field $\boldsymbol{B}$. Let θ be the angle of inclination of the axis of the atomic dipole with the direction of the applied field $\boldsymbol{B}$. Then magnetic potential energy of the atomic dipole is

$$U = -mB\cos\theta$$

Now, on classical statistics, the number of atoms making an angle between θ and $\theta + d\theta$ is

$$dn = Ce^{mB\cos\theta/kT}\sin\theta\, d\theta$$

where k is Boltzmann's constant and T is the absolute temperature.

Put $mB/kT = \alpha$. Then

$$dn = Ce^{\alpha\cos\theta}\sin\theta\, d\theta \qquad ...(1)$$

Hence the total number of atomic magnets in unit volume of the paramagnetic material

$$n = \int_0^\pi dn = \int_o^\pi Ce^{\alpha\cos\theta}\sin\theta\, d\theta \qquad ...(2)$$

Put $\cos\theta = x$. Then $-\sin\theta\, d\theta = dx$.

$$n = \int_{+1}^{-1} -Ce^{\alpha x}\, dx = C\int_{-1}^{+1} e^{\alpha x}\, dx$$

$$\therefore \quad C = \frac{n\alpha}{e^{\alpha} - e^{-\alpha}} \qquad ...(3)$$

The component of each dipole moment parallel to B in $m\cos\theta$. The total magnetic moment of all the n atoms contained in unit volume of the gas is the magnetisation M. It is given by

$$M = \int_0^\pi m\cos\theta\, dn = \int_0^\pi m\cos\theta\, Ce^{\alpha\cos\theta}\sin\theta\, d\theta \qquad ...(4)$$

Put $\cos\theta = x$. Then, $-\sin\theta\, d\theta = dx$. Therefore, we get

$$M = \int_{+1}^{-1} -mxCe^{\alpha x}\, dx = Cm\int_{-1}^{+1} xe^{\alpha x}\, dx$$

Evaluating this integral and substituting the value of C from (3), we get

$$M = mn\left[\frac{e^{\alpha}+e^{-\alpha}}{e^{\alpha}-e^{-\alpha}}-\frac{1}{\alpha}\right]$$

$$= mn\left[\coth\alpha-\frac{1}{\alpha}\right]$$

$$= mn\, L\,(\alpha) \qquad \text{...(5)}$$

where $L\,(\alpha) = \left[\coth\alpha-\frac{1}{\alpha}\right]$ is called the *Langevin function.*

The variation of M with α is shown in Fig. 47.2.

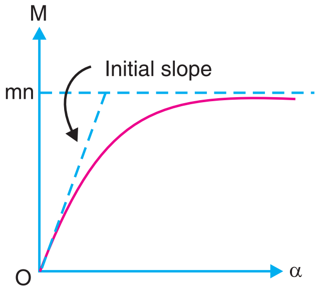


Fig. 47.2

Case (*i*): At low temperatures or large applied field, $L\,(\alpha) \to 1$.

Hence, magnetisation M in this case will be

$$M = mn \qquad \text{...(6)}$$

So saturation is reached when all the atomic dipoles are parallel to B.

Case (*ii*) : Under normal conditions α is very small. Then,

$$L(\alpha) = \coth\alpha-\frac{1}{\alpha} \approx \frac{\alpha}{3} \qquad \text{...(7)}$$

$$\therefore \quad M = mn\frac{\alpha}{3} = \frac{nm^2B}{3kT} = \frac{nm^2\mu_0 H}{3kT} \qquad \text{...(8)}$$

$$\chi = \frac{M}{H} = \frac{\mu_0 nm^2}{3kT} = \frac{C}{T} \qquad \text{...(9)}$$

where $C = \mu_0 nm^2/3k$ is called the Curie constant.

Failure of Langevin Theory. (*i*) Langevin's theory was unable to explain a more complicated dependence of susceptibility upon temperature exhibited by several paramagnetics such as highly compressed and cooled gases, very concentrated solutions of salts, etc.

(*ii*) Langevin's theory could not account for the intimate relation between para-and ferro-magnetism.

Weiss Modification: Langevin's theory applies strictly only to gases, where the molecules are sufficiently far apart for their mutual interactions to be negligible. In liquids and solids such interactions may be large, and many substances obey the modified *Curie-Weiss law*

$$\chi = C/(T-\theta) \qquad \text{...(10)}$$

θ is called the *Curie temperature* and is characteristic of the substance. Eq. (10) holds only at temperatures where $T > |\,\theta\,|$. Eq. (10) is of the same form as Eq. (9), except that the origin of temperature is shifted from 0 to θ.

47.4 Ferromagnetism

Ferromagnetic substances are *very strongly magnetic.* The best-known examples of ferromagnets are the transition metals *Fe*, *Co*, and *Ni*. A ferromagnet has a spontaneous magnetic moment–a magnetic moment even in zero applied field. The atoms (or molecules) of ferromagnetic materials have a net intrinsic magnetic dipole moment which is primarily due to the spin of the electrons. The

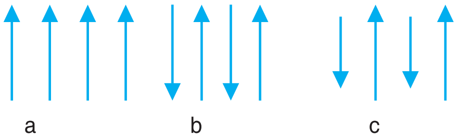
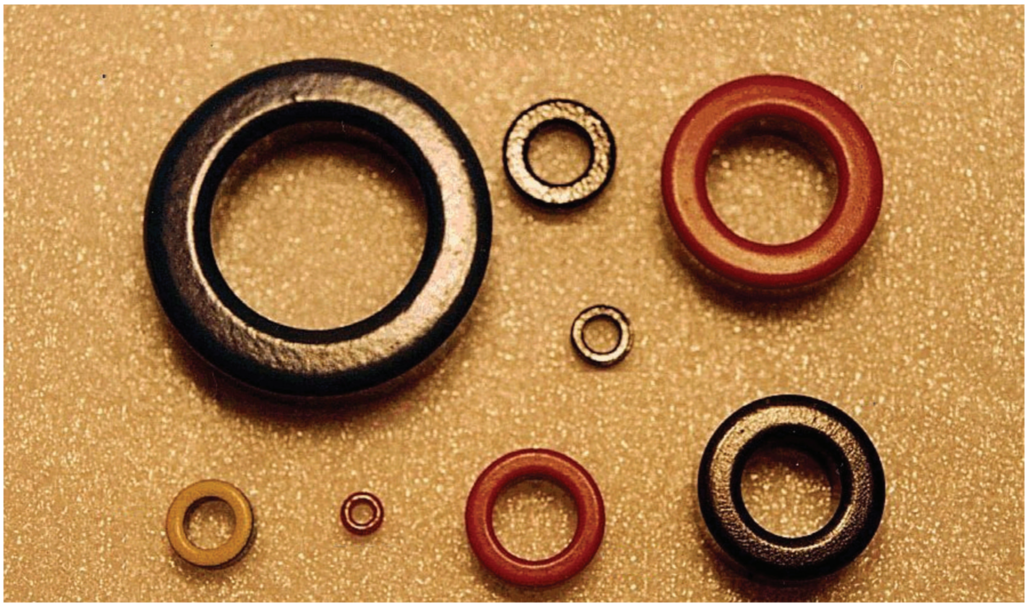


Fig. 47.3

interaction between the neighbouring atomic magnetic dipoles is very strong. It is called *spin exchange interaction* and is present even in the absence of an external magnetic field. It turns out that the energy of two neighbouring atomic magnets due to this interaction is the least when their magnetic moments are parallel. The neighbouring magnetic moments are therefore strongly constrained to take parallel orientation (Fig. 47.3*a*). This effect of the exchange interaction to align the neighbouring magnetic moments parallel to one another spreads over a small finite volume of the bulk.



Ferromagnetic Materials

This small (1-0.1 mm across) volume of the bulk is called a *domain.* All magnetic moments within a domain will point in the same direction, resulting in a large magnetic moment. Thus the bulk material consists of many domains. The domains are oriented in different directions. The total magnetic moment of a sample of the substance is the vector sum of the magnetic moments of the component domains.

In an unmagnetised piece of ferromagnetic material, the magnetic moments of the domains themselves are not aligned. When an external field is applied, those domains that are aligned with the field increase in size at the expense of the others. In a very strong field all the domains are lined up in the direction of the field and provide the high observed magnetisation.

Antiferromagnetism and Ferrimagnetism. The only type of magnetic order which has been considered thus far is ferromagnetism, in which, in the fully magnetised state, all the dipoles are aligned in exactly the same direction (Fig. 47.3*a*). There are, however, substances which show different types of magnetic order. In *antiferromagnetic* materials such as Cr and MnO, the dipoles have equal moments, but adjacent dipoles point in opposite directions (Fig. 47.3*b*). Thus the moments balance each other, resulting in a zero net magnetisation. In *ferrimagnetic materials* (also called *ferrites*) such as $Mn\ Fe_2O_4$, the magnetic moments of adjacent ions are antiparallel and of unequal strength (Fig. 47.3*c*). So there is a finite net magnetisation. By suitable choice of rare-earth ions in the ferrite lattices it is possible to design ferrimagnetic substances with specific magnetizations for use in electronic components.

47.5 Soft and Hard Magnetic Materials

The process of magnetisation of a ferromagnetic material consists of moving the domain walls so that favourably oriented domains grow and unfavourably oriented domains shrink.

- If the domain walls are easy to move, the coercive field is low. It is easy to magnetise the material. Such a material is called a *soft magnetic material.*
- If it is difficult to move the domain walls, the coercive field is large and the material is magnetically *hard.*

Fig. 47.4 shows *magnetisation curves for soft and hard magnetic materials.*

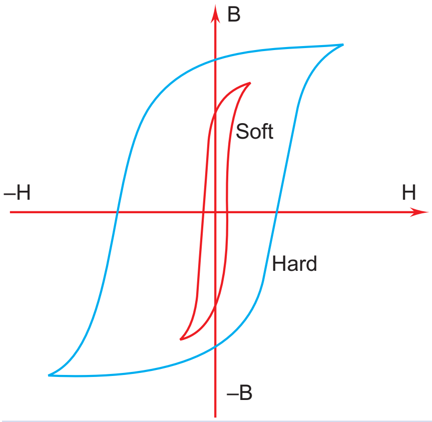
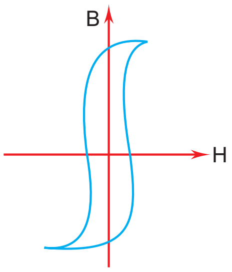


Fig. 47.4

Soft Magnetic Materials

Fig. 47.5 shows the nature of hysteresis loop of soft magnetic material (soft iron).

- Soft magnetic materials have low hysteresis loss due to small hysteresis loop area.
- In these materials, the domain wall movement is relatively easier. Even for small changes in the magnetizing field, magnetization changes by large amounts.
- The coercivity and retentivity are small. Hence these materials can be easily magnetized and demagnetized.
- Soft magnetic materials are used in applications requiring frequent reversals of the directions of magnetisation such as cores of transformers.



Soft iron

Fig. 47.5

- In soft magnetic materials, the hysterisis losses must be kept down to a minimum. When the magnetic induction is large for a small applied field, the loop area is small and the hysteresis loss is reduced. The key factor in the design of a soft magnet is then to have easily moving domain walls. Soft magnetic materials should be free of impurities and inclusions.

Usually there are easy and hard magnetisation directions in a crystal. Fig. 47.6 shows magnetisation curves for single crystals of iron. Iron magnetises more easily along (1 0 0) than along (1 1 1). (1 1 1) is the hard direction for iron. This property can be exploited to reduce the area under the hysteresis loop by manufacturing materials with a preferred orientation of grains.

- The other source of energy loss in soft magnetic materials is the eddy current loss (changing magnetic flux in a medium induces an emf). The induced emf is proportional to the rate of change of flux and hence to the frequency of the alternating current. The induced emf sets up eddy current. The power loss due to these is equal to V^2/R. Here, V is the induced emf and R is the resistance of the medium. Eddy current losses can be minimised by increasing the resistivity of the medium.

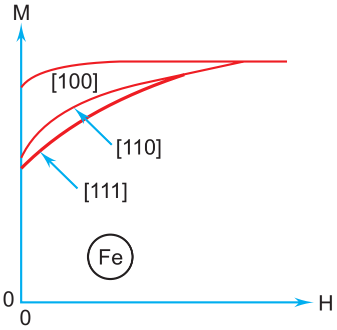


Fig. 47.6

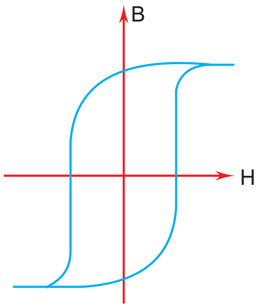
Iron, which used to be the material for transformer cores, is now almost entirely replaced by Fe–Si alloys, which has substantially higher resistivity than iron. Fe–Si alloys are suitable for operation at power frequencies of 50–60 Hz. At microwave frequencies, ferrites (48% $MnO \cdot Fe_2O_3$, 52% $ZnO \cdot Fe_2O_3$; 52% $ZnO.Fe_2O_3$, 36% $NiO \cdot Fe_2O_3$, 64% $ZnO \cdot Fe_2O_3$) and garnets ($3Y_2O_3 \cdot 5Fe_2O_3$) are preferred.

Hard Magnetic Materials

Fig. 47.7 shows the nature of hysteresis loop of hard magnetic material (steel).

- Hard magnetic materials have large hysteresis loss due to large hysteresis loop area.
- In these materials, the domain wall movement is difficult because of presence of impurities and crystal imperfections and it is irreversible in nature.
- The coercivity and retentivity are large.

Hence, these materials cannot be easily magnetized and demagnetized.

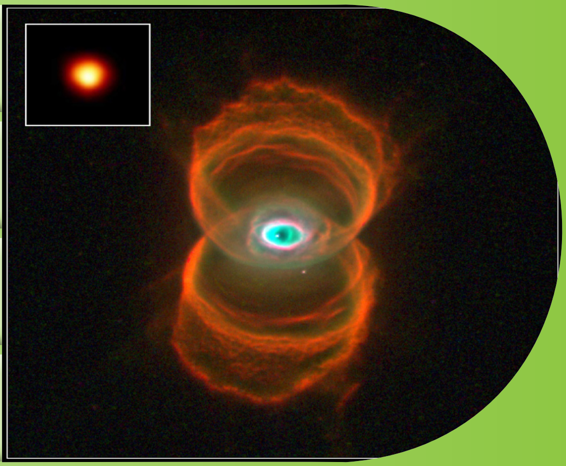


Steel

Fig. 47.7

Hard magnetic materials are used to produce permanent magnets. Hysteresis losses are of no significance here as no repeated reversals of magnetisation is involved in a permanent magnet. The permanent magnets must have high residual induction B, and large coercive field H_c. The area of the hysteresis loop between B_r and H_c represents the energy required to demagnetise a permanent magnet. The maximum value of this area (= B_rH_r), called the *energy product,* must be as large as possible for permanent magnets. High carbon steels and other low alloy tungsten and chromium steels are used for making permanent magnets.

CHAPTER 48



MAGNETISM

AT A GLANCE

48.1 Weiss Theory of Ferromagnetism
48.2 Nuclear Magnetic Resonance

48.1 Weiss Theory of Ferromagnetism

According to Weiss, the atomic magnets of a ferromagnetic substance are grouped into certain regions or *domains*. When the substance is in the unmagnetised condition, the domains form closed chains with no free poles. When the substance is magnetized, the chains break up and the domains gradually set themselves with their magnetic axes all pointing in the field direction. Thus ferromagnetism is a *crystal phenomenon*.

Weiss assumed that a molecular magnetic field exists at the position of every atom or molecule. This field arises due to the interaction of all neighbouring molecules. The molecular field is proportional to the magnetisation vector I.

$\therefore$ molecular field $B_i = \beta I$

Here, β = molecular field coefficient.

The effective field strength B_e may be regarded as the vector sum of external applied field strength B and the internal molecular field strength B_i.

Hence

$$B_e = B + B_i = B + \beta I \qquad \text{...(1)}$$

Consider 1 gram mole of the substance. Let

ρ = density of the substance

M = molecular weight

σ = gram molecular magnetic moment

σ_0 = saturation value of gram molecular magnetic moment.

Then

$$I = \frac{\sigma}{(M/\rho)} \quad \text{and} \quad I_s = \frac{\sigma_0}{(M/\rho)}$$

$$\therefore \qquad \frac{I}{I_s} = \frac{\sigma}{\sigma_0} \qquad \text{...(2)}$$

The domains will obey the general theory of paramagnetism by Langevin.

$$\therefore \qquad \frac{I}{I_s} = \frac{\sigma}{\sigma_0} = \left(\coth\alpha - \frac{1}{\alpha}\right) \qquad \text{...(3)}$$

Here, $$\alpha = \frac{\mu B_e}{kT} = \frac{\mu(B+\beta I)}{kT}$$

μ = dipole moment of individual atom (molecule)

k = Boltzmann constant

T = absolute temperature of substance.

When external field is zero,

$$B_e = \beta I = \beta\frac{\sigma\rho}{M}$$

Now, $$\alpha = \frac{\mu\beta\sigma\rho}{kTM} = \frac{\sigma_0}{N} \times \frac{\beta\sigma\rho}{kTM}$$

Here, N = Avogadro's number

or $$\alpha = \frac{\sigma_0 \beta\sigma\rho}{RTM}$$

or $$\frac{\sigma}{\sigma_0} = \frac{RTM}{\sigma_0^2\beta\rho}\alpha \qquad \text{...(4)}$$

Eqs. (3) and (4) simultaneously determine the condition of spontaneous magnetisation. Fig. 48.1 shows a graph drawn between σ / σ_0 and α corresponding to Eqs. (3) and (4). Curve 1 is Langevin curve corresponding to Eqs. (3) while straight line 2 is corresponding to Eqs. (4). The two curves meet at 0 and A. The solution $\frac{\sigma}{\sigma_0} = 0$ is not true. Hence $\frac{\sigma}{\sigma_0} = AD$ is the true solution. It is obvious from the graph that A represents the stable state of spontaneous magnetisation. If the molecules in a domain assume state C, then the local magnetisation is less than the corresponding equilibrium state A. Now the magnetisation and the value of α will in consequence increase until the state AD is reached. On the other hand, if the molecules in a domain assume state B, then the local magnetisation is more than the equilibrium value. Now the magnetisation and the value of α will tend to decrease until the state AD is reached.

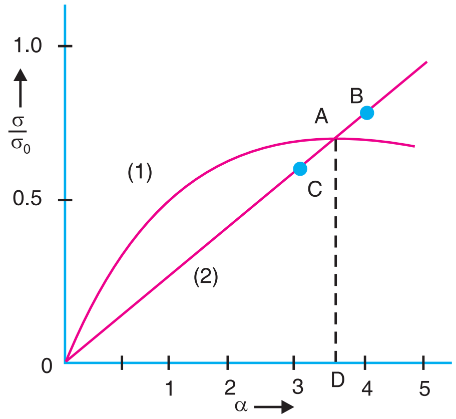


Fig. 48.1

We know that the slope of the tangent at the origin of the Langevin curve is $\frac{1}{3}$ since $\frac{\sigma}{\sigma_0} = \frac{\alpha}{3}$, when α is small. Hence, the condition for *stable spontaneous magnetisation* is given by

$$\frac{RTM}{\sigma_0^2\beta\rho} < \frac{1}{3} \quad \text{or} \quad T < \frac{\sigma_0^2\beta\rho}{3RM}$$

But $\frac{\sigma_0^2 \beta \rho}{3RM} = \theta$, the Curie point. Hence $T < \theta$.

Hence below the Curie point θ, in the absence of the external field, the domains are spontaneously magnetised to a degree depending upon temperature, approaching the saturation value as the temperature approaches absolute zero. Above the Curie point θ, spontaneous magnetisation no longer occurs, the ferromagnetic properties disappear and the substance becomes paramagnetic. Therefore, Curie temperature is transition temperature. Above Curie point, the substance obeys Curie-Weiss law.

The chief drawbacks of Weiss theory are that (*i*) it could not explain why and how internal field between the molecules of ferromagnetic materials possess such large values, and (*ii*) why the linear relationship expressed by Curie-Weiss law breaks near Curie point.

48.2 Nuclear Magnetic Resonance

A nucleus with a nonzero spin acts like a tiny magnet. When placed in a magnetic field, the magnetic moment vector precesses round the field direction. This is called Larmor precession. This precession is quantised. If this precessing nucleus is placed in the path of a beam of electromagnetic radiation there will be coherent interaction between the particle and the radiation provided the frequency of the radiation is equal to that of precession. The phenomenon is thus one of *resonance*. In the case of nuclei, the phenomenon is called *nuclear magnetic resonance* (N. M. R).

Nuclear Magnetic Resonance (NMR).

Expression for Resonance frequency

If a nucleus of magnetic moment μ is placed in a uniform magnetic field B_0 in the z-direction, then its energy is

$$U = -\mu_z B_0 \qquad \text{...(1)}$$

The magnetic moment μ_z is quantized according to the usual rule

$$\mu_z = g m_I \frac{e\hbar}{2m_p} \qquad m_I = I, I-1, ..., 1-I, -I \qquad \text{...(2)}$$

where g is the g-factor for the nucleus. Hence the energy levels of the nucleus in the magnetic field are

$$U = -g m_I \frac{e\hbar}{2m_p} B_0 \qquad m_I = I, I-1, ..., 1-I, -I \qquad \text{...(3)}$$

The energy difference between adjacent energy levels is

$$\Delta U = -g e \hbar B_0 / 2(m_p) \qquad \text{...(4)}$$

The corresponding transition frequency is

$$\nu = \frac{\Delta U}{h} = \frac{g e B_0}{4\pi m_p}.$$

NMR Spectrometer. Fig. 48.2 shows a schematic arrangement for observing the nuclear magnetic resonance, and is based on the energy absorption method.

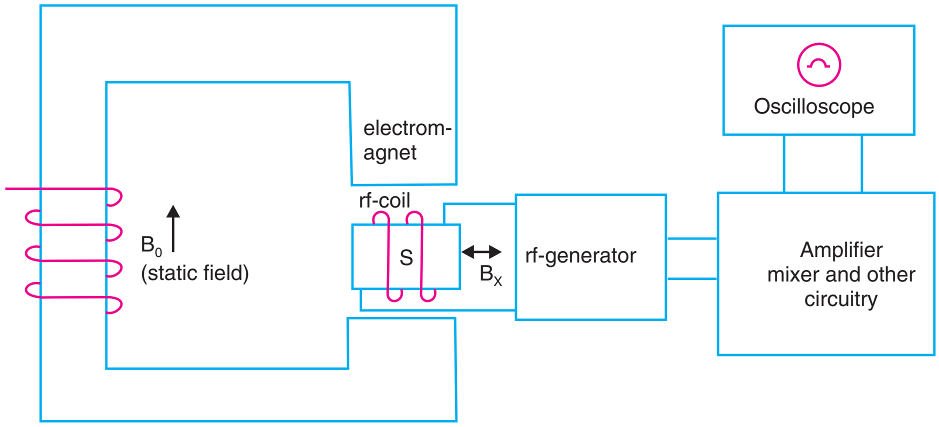


Fig. 48.2

The specimen *S*, about 1 c.c. of the material being investigated, is placed between the poles of an electromagnet. This magnet produces the field B_0. An r-f coil surrounding the specimen is carefully positioned to produce a second field B_x perpendicular to B_0. An r-f generator not only serves to drive the coil but also supplies a signal to auxiliary circuitry which measures the r-f power absorbed by the specimen. To trace out the absorption line, an auxiliary low-frequency oscillator supplies power to secondary coils wound on the main magnet core. These coils permit the main field B_0 to be varied by a small amount on both sides of the resonant field. Associated electronic circuitry displays on an oscilloscope, a signal proportional to the energy absorbed as a function of field. Such a signal is sketched in Fig. 48.3.

BRUKER
600
UltraShield™

NMR Spectrometer.

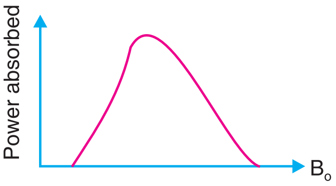


Fig. 48.3

NMR finds a number of practical applications. It can be used for the accurate determination of nuclear moments. In a sensitive form of magnetometer NMR can be used for measuring magnetic fields. In medicine NMR tomography has been developed. In this technique images of tissues are produced by magnetic resonance. Whole plants, small animals or parts of bigger animals are placed between the pole pieces of a suitably designed magnet. Using surface coils of wire radio frequency radiation is directed towards the part under study. Spectra obtained in this way are used for medical diagnosis. This method is supposed to be superior to diagnosis by X-rays because the potential damage to living tissue by X-rays is not present in the case of radio frequency radiation.

The main application of NMR is in its use for chemical analysis and determination of molecular structures.

CHAPTER

49

QUANTUM THEORY OF PARAMAGNETISM

AT A GLANCE

49.1	Quantum Theory of Paramagnetism
49.2	Cooling by Adiabatic Demagnetization of a Paramagnetic salt

49.1 Quantum Theory of Paramagnetism

The magnetic moment of an atom or ion in free space is given by

$$\vec{\mu} = \gamma\hbar\mathbf{J} = -g\,\mu_B\mathbf{J} \qquad ...(1)$$

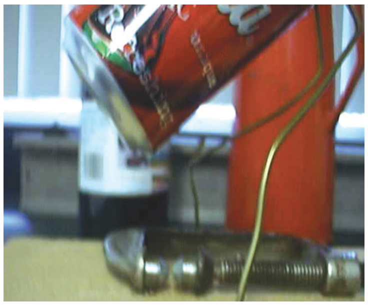
Here, total angular momentum $\hbar\mathbf{J}$ = orbital angular momentum $\hbar\mathbf{L}$ + spin angular momentum $\hbar\mathbf{S}$.

The ratio of the magnitude of $\vec{\mu}$ to that of **J** is called the *gyromagnetic ratio* (γ) or *magnetogyric ratio.* μ_B is the Bohr magneton. It is defined as $e\hbar/2m$.

For a free atom the g factor is given by the Lande equation

$$g = 1 + \frac{J(J+1) + S(S+1) - L(L+1)}{2J(J+1)} \qquad ...(2)$$

For an electron spin $g = 2.00$.



Low Cost Paramagnetism Apparatus.

The energy levels of the system in a magnetic field are

$$U = -\vec{\mu} \cdot \mathbf{B} = m_j g\mu_B \mathbf{B} \qquad \text{...(3)}$$

Here, m_J is the azimuthal quantum number.

$$m_J = J, (J-1), \ldots -(J-1), -J$$

For a single spin with no orbital moment, we have $m_J = \pm\frac{1}{2}$ and $g = 2$.

$$\therefore \qquad U = \pm\mu_B \mathbf{B}. \qquad \text{...(4)}$$

Thus an energy difference exists between spin magnetic dipole moments parallel and anti-parallel to the field. The magnitude of energy difference is $2\mu_B \mathbf{B}$. This splitting is shown in Fig. 49.1.

Let us suppose that in unit volume of the paramagnetic material there is a total of N atoms. The spin moments of N_1 atoms are parallel to the field and N_2 spin moments are antiparallel.

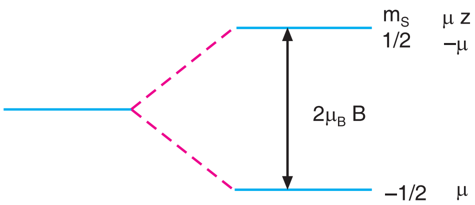


Fig. 49.1

If a system has only two levels the equilibrium populations are,

$$\text{with } \tau = k_B T,$$

$$\frac{N_1}{N} = \frac{\exp(\mu B/\tau)}{\exp(\mu B/\tau) + \exp(-\mu B/\tau)}; \qquad \text{...(5)}$$

$$\frac{N_2}{N} = \frac{\exp(-\mu B/\tau)}{\exp(\mu B/\tau) + \exp(-\mu B/\tau)}. \qquad \text{...(6)}$$

Here N_1 and N_2 are the populations of the lower and upper levels, and $N = N_1 + N_2$ is the total number of atoms. The fractional populations are plotted in Fig. 49.2.

The projection of the magnetic moment of the upper state along the field directions is – μ and of the lower state is μ. The resultant magnetization for N atoms per unit volume is

$$M = (N_1 - N_2)\mu = N\mu . \frac{e^x - e^{-x}}{e^x + e^{-x}} = N\mu \tanh x \qquad \text{...(7)}$$

$$\text{Here,} \qquad x \equiv \mu B/k_B T \qquad \text{...(8)}$$

For $x << 1$, tan h $x \cong x$, and we have

$$M \cong N\mu\,(\mu B/k_B T) \qquad \text{...(9)}$$

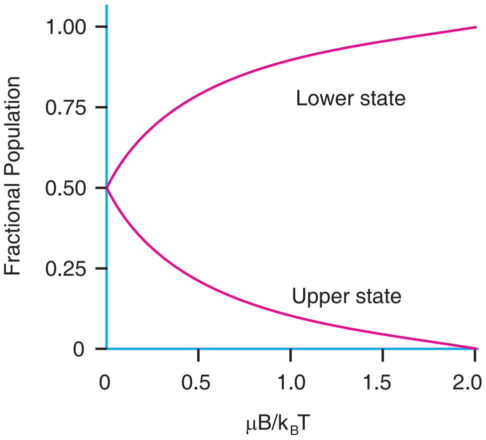


Fig. 49.2

In a magnetic field an atom with angular momentum quantum number J has $(2J + 1)$ equally spaced energy levels. The Magnetization is given by

$$M = NgJ\mu_B B_J(x) \qquad \text{...(10)}$$

$$\text{Here,} \qquad x \equiv gJ\mu_B B/k_B T \qquad \text{...(11)}$$

Eq. (10) is called the *Curi-Brillouin law.*

$B_J(x)$ is called the *Brillouin function* defined as

$$B_j(x) = \frac{2J+1}{2J}\operatorname{ctnh}\left(\frac{(2J+1)x}{2J}\right) - \frac{1}{2J}\operatorname{ctnh}\left(\frac{x}{2J}\right) \qquad \text{...(12)}$$

Equation (7) is a special case of Eq. (12) for $J = \frac{1}{2}$.

Fig. 49.3 shows the plot of magnetic moment versus B/T for spherical samples of (**I**) potassium chromium alum, (**II**) ferric ammonium alum, and (**III**) gadolinium sulphate octahydrate.

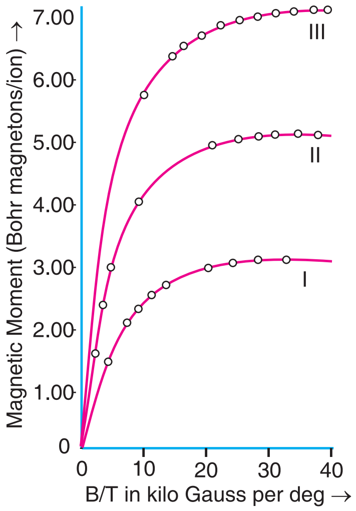


Fig. 49.3

For $x << 1$, we have

$$\text{etnh} = \frac{1}{x} + \frac{x}{3} - \frac{x^3}{45} + \ldots, \qquad \ldots(13)$$

The susceptibility per unit volume is

$$\chi = \frac{M}{B} \cong \frac{NJ(J+1)g^2\mu_B^2}{3k_B T} = \frac{Np^2\mu_B^2}{3k_B T} = \frac{C}{T} \qquad \ldots(14)$$

Here p is called the *effective number of Bohr Magnetons*. It is defined as

$$p \equiv g\,[J\,(J+1)]^{1/2} \qquad \ldots(15)$$

Eq. (14) is called the *Curie law. C* is called the *Curie constant.*

Rare Earth Ions

Equation (14) is successfully employed to predict the values of susceptibility for various paramagnetic crystals particularly rare earth ions. The value of J is determined by applying the Hund's rules.

Experiments on rare-earth ions in crystals show that they obey the Curie law, with an effective number of magnetons in agreement with the theory of spin-orbit interaction. In these ions, therefore, the angular momenta **L** and **S** are strongly coupled, and the moment of the ion can respond freely to the external field.

In these ions–from La to Lu in the periodic table–the 4*f* shell is incompletely filled. The outer 5*p* shell is completely filled, while the 5*d* and 6*s* shells which are still further out are stripped of their electrons to form the ionic crystal (Fig. 49.4). Thus the only incomplete shell is the 4*f* shell. This is the one in which the magnetic behaviour occurs. Electrons in this shell lie deep within the ion, screened by the outer 5*p* and 5*d* shells. So they are not appreciably affected by other ions in the crystal. Magnetically their behaviour is much like that of a free ion.

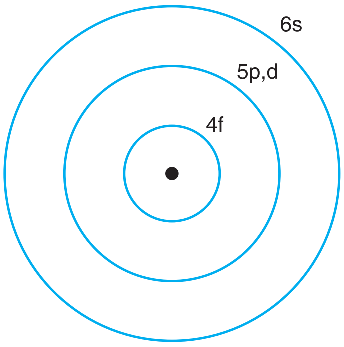


Fig. 49.4

Iron-group ions

Equation (14) is unable to account for the experimental observations of susceptibility for the ions of the iron group. The reason for this is that for the iron group, the ground state of the ions in a crystal is very different from that of the free ion. This is because, the partly filled electron shell, which is responsible for producing the magnetic moment, is the outermost (3*d*) shell of these ions. It is, therefore, very strongly influenced by the crystal field. The interaction is so strong that the orbital angular momentum of the electrons in the 3*d* shell is very often reduced to zero. This does not mean that the electrons no longer encricle the nucleus, but

Quenching is an Important Phenomenon in Superconducting Magnets.

rather that they will be travelling around it as often in one sense as in the other sense. This effect when L averages to zero is termed *quenching* of the orbital angular momentum and it means that when the ion is in the crystal, there is no orbital magnetic moment. In quenching, the crystal field hardly interacts with the spin and hence, the spin magnetic moment still exists. Therefore, while calculating susceptibility on the basis of Eq. (14), we should replace J by S. When this is done, quite good agreement is obtained between the theoretical and experimental Curie constant.

49.2 Cooling by Adiabatic Demagnetization of a Paramagnetic Salt

Adiabatic demagnetization of a large paramagnetic sample can be used for attaining temperatures much below 1K. By its use temperatures of 10^{-3} K and lower have been reached. The principle of the method is that at a fixed temperature the entropy of a system of magnetic moments is lowered by the application of a magnetic field.

The entropy is a measure of the disorder of a system. Therefore, the greater the disorder, the higher is the entropy. In the magnetic field the magnetic moments will be partly lined up *i.e.*, partly ordered, so that the entropy is lowered by the magnetic field. The entropy is also lowered if the temperature is lowered, as more of the magnetic moments line up.

Suppose the magnetic field is removed without changing the entropy of the spin system. Then, the order of the spin system will look like a lower temperature than the same degree of order in the presence of magnetic field. When the specimen is demagnetized adiabatically, entropy can flow into the spin system only from the lattice vibrations (Fig. 49.5). At very low temperatures, the entropy of lattice vibrations is usually negligible. Thus, the entropy of the spin system will be essentially constant during adiabatic demagnetization of the specimen. It is important to mention that magnetic cooling is a one-shot operation, and it is not cyclic.

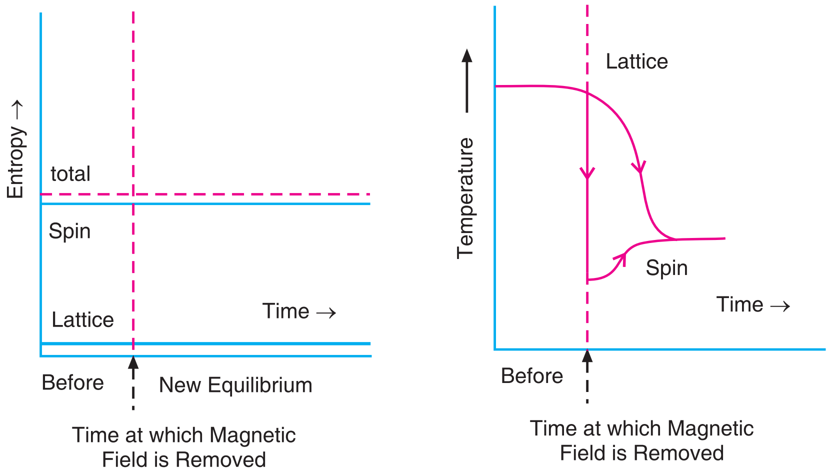


Fig. 49.5

Let us find an expression for the spin entropy of a system of N ions, each of spin S, at a temperature sufficiently high that the spin system is entirely disordered. That is, T is supposed to be much higher than some temperature θ which characterizes the energy of the interactions ($E_{int} = k_B \theta$) tending to orient the spins preferentially. The entropy σ of a system of G accessible states is $\sigma = k_B$ 1n G. At a temperature so high that all of the $(2S + 1)$ states of each ion are nearly equally populated, G is the number of ways of arranging N spins in $(2S + 1)$ states. Thus, we have

$$G = (2S + 1)^N.$$

The spin entropy

$$\sigma_S = k_B \, 1\text{n}\, (2S + 1)^N = N k_B \, 1\text{n}\, (2S + 1) \qquad ...(1)$$

This spin entropy is reduced by a magnetic field if the lower levels gain in population when the field separates the $2S + 1$ states in energy. Fig 49.6 shows the steps carried out in the cooling process. Magnetic field is applied at temperature T_1 with the specimen in good thermal contact with the surroundings, giving the isothermal path *ab*. The specimen is then insulated ($\Delta\sigma = 0$) and the magnetic field is removed. Thus the specimen follows the constant entropy path *bc*, ending up at temperature T_2. At temperature T_1 the thermal contact is provided by helium gas and by removing the gas with a pump the thermal contact is broken.

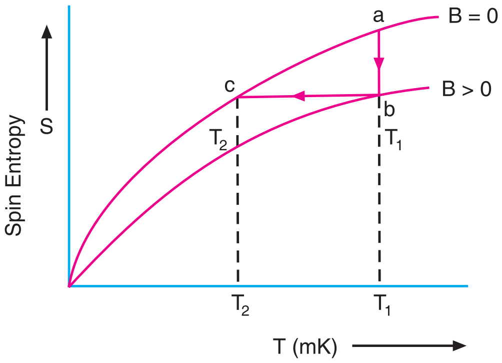


Fig. 49.6

EXERCISE

1. Give an account of the quantum theory of paramagnetism and derive an expression for the susceptibility. How does this theory account for the experimentally observed susceptibility of the rare earth ions? (*Raj.,* 1978)

CHAPTER 50

LATTICE VIBRATIONS AND PHONONS

AT A GLANCE

50.1	Introduction
50.2	Vibrations of Monatomic Linear Lattice
50.3	Vibrational Modes of a Finite One-dimensional Lattice of Identical Atoms
50.4	Vibrational Modes of a Diatomic Linear Lattice
50.5	Quantisation of Lattice Vibrations
50.6	Inelastic Scattering of Neutrons by Phonons

50.1 Introduction

Our discussion so far in the previous chapters have been concerned with systems involving electrons. We now switch emphasis in this chapter to some properties of solids not dependent on the presence of mobile electrons but dependent rather on the motions of the much more massive ion cores. While these atomic structures do not generally diffuse throughout the crystal, they do vibrate about the lattice sites and create collective wave motions that are called *phonons.*

A lattice is usually regarded as an array of atoms connected with each other by elastic springs. With this model for the lattice in mind, it is quite easy to understand that the motion of the adjacent atoms (excited, say, by dynamical external forces) would be coupled, or in other words, we can say that the motion of every atom would be shared by all the atoms in the crystal and the crystal would thus vibrate as a whole. Our problem is to investigate the characteristics of this elastic vibrational motion of the crystal lattices.

50.2 Vibrations of Monatomic Linear Lattice

A monatomic linear crystal lattice is an array of identical, equidistant atoms (or ions), each of mass m, with interatomic separation a. The atoms are presumed to be connected by springs and the displacements of atoms are within the Hooke's law limits. Further, only nearest neighbour interactions are considered. When an elastic wave propagates through the crystal, the atoms are set in vibrations either parallel or perpendicular to the direction of the wave vector. We shall consider longitudinal elastic vibrations of the atoms.

Let us consider an infinitely long linear chain of identical atoms (Fig. 50.1). The open circles represent the equilibrium positions and the black dots the displaced atoms.

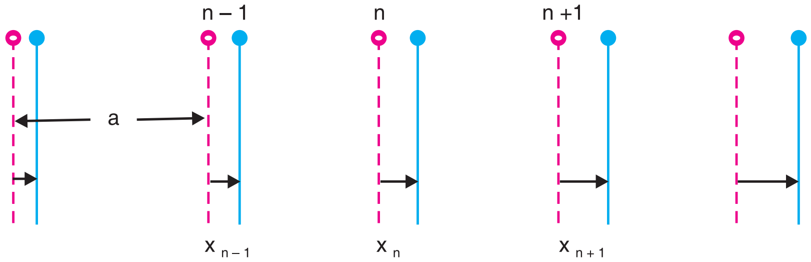


Fig. 50.1

Let x_n = displacement of nth atom from equilibrium position

x_{n-1} = displacement of $(n-1)$th atom from equilibrium position

x_{n+1} = displacement of $(n+1)$th atom from equilibrium position

F_n = force on nth atom, and

f = force constant for the nearest-neighbour interaction.

According to Hooke's law, the force on an atom due to the displacement of its (nearest) neighbour is proportional to the difference of their displacements. So the net force on the nth atom is

$$\begin{aligned} F_n &= f(x_{n+1}-x_n)-f(x_n-x_{n-1}) \\ &= f(x_{n+1}+x_{n-1}-2x_n) \end{aligned}$$

The equation of motion of the nth atom is

$$m\ddot{x}_n = f(x_{n+1}+x_{n-1}-2x_n). \quad ...(1)$$

Let a solution to this equation be of the form of a travelling wave :

$$x_n = e^{-i\omega(t-na/v)} \quad ...(2)$$

Here, na = the equilibrium position (x-co-ordinate) of the nth particle relative to the origin, and

v = the velocity of wave propagation and $v = \nu\lambda$

ω = the angular frequency = $2\pi\nu$.

Now, the wave vector $K = \dfrac{2\pi}{\lambda} = \dfrac{\omega}{v}$.

Therefore, Eq. (2) can be written as

$$x_n = e^{-i(\omega t - Kna)} \quad ...(3)$$

$$\therefore \quad x_{n+1} = e^{-i\{\omega t - K(n+1)a\}} = e^{iKa}x_n$$

and $$x_{n-1} = e^{-i\{\omega t - K(n-1)a\}} = e^{-iKa}x_n.$$

Differentiating Eq. (3) twice with respect to t, we get

$$\ddot{x}_n = \frac{d^2x_n}{dt^2} = (-i\omega)^2 e^{-i(\omega t - Kna)} = -\omega^2 x_n.$$

Substituting for $\ddot{x}_n, x_{n+1}, x_{n-1}$ and x_n in Eq. (1), and dividing by x_n, we get

$$\begin{aligned} -m\omega^2 &= f(e^{iKa}+e^{-iKa}-2) \\ &= f(e^{iKa/2}-e^{-iKa/2})^2. \end{aligned}$$

Now, $$\sin\frac{Ka}{2} = \frac{e^{iKa/2} - e^{-iKa/2}}{2i}$$

or $$\sin^2\frac{Ka}{2} = -\frac{1}{4}\left(e^{iKa/2} - e^{-iKa/2}\right)^2.$$

∴ $$-m\omega^2 = -4f\sin^2\frac{Ka}{2}$$

or $$\omega^2 = \frac{4f}{m}\sin^2\frac{Ka}{2}$$

or $$\omega = 2\sqrt{\frac{f}{m}}\left|\sin\frac{Ka}{2}\right|. \qquad \text{...(4)}$$

The absolute value signs occur because a frequency is always a positive quantity by definition. Fig. 50.2 gives the plot of ω versus K. It shows that frequency is a periodic function of K, with period $2\pi/a$. The dashed lines correspond to a continuous string for which $\omega \propto K$.

Eq. (4) is called the dispersion relation.

Results— We draw the following results from the above discussion :

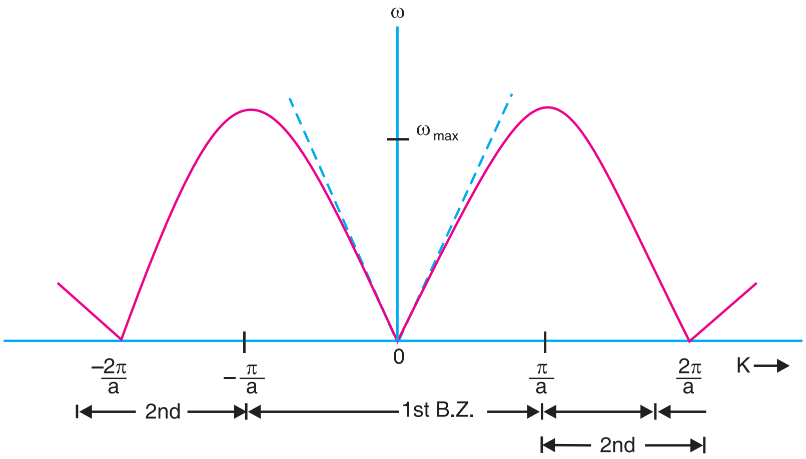


Fig. 50.2

1. At low frequencies ($K \to 0$, the long wavelength limit)

$$\sin\frac{Ka}{2} \to \frac{Ka}{2}$$

∴ $$\omega = Ka\sqrt{\frac{f}{m}} \qquad \text{...(5)}$$

In this long wavelength approximation the phase velocity will be essentially constant, since

$$v_p = \frac{\omega}{K} = a\sqrt{\frac{f}{m}} = v_0 \qquad \text{...(6)}$$

Group velocity

$$v_g = \frac{d\omega}{dK} = a\sqrt{\frac{f}{m}} = v_0 \qquad \text{...(7)}$$

∴ $$v_g = v_p = \text{constant}$$

For very long wavelengths then, the dispersion effects are negligible and the medium acts like a continuous and homogeneous elastic medium. That is what happens in a continuous string. Thus, *for long wavelengths*, the periodicity of the lattice is of little importance.

For shorter wavelengths the approximation of Eq. (5) is no longer valid and waves of different frequencies travel at different velocities. The dispersion arises because of the discrete character of this chain of particles.

2. Another important result is obtained by comparing the solution Eq. (3) with another in which K is replaced by K_m where

$$K_m = K + \frac{2\pi m}{a} \text{ with } m = \pm 1, \pm 2, \ldots\ldots$$

We have for the displacement of the nth atom

$$\begin{aligned} x_n &= e^{-i(\omega t - K_m na)} \\ &= e^{-i\{\omega t - (K + 2\pi m/a)na\}} \\ &= e^{-i(\omega t - Kna)} \qquad [\because e^{2\pi im} = 1] \end{aligned}$$

From Eq. (4), we have

$$\begin{aligned} \omega &= \sqrt{\frac{4f}{m}} \left| \sin \frac{1}{2} K_m a \right| \\ &= \sqrt{\frac{4f}{m}} \left| \sin \frac{1}{2} a \left(K + \frac{2\pi m}{a} \right) \right| \\ &= \sqrt{\frac{4f}{m}} \left| \sin \frac{1}{2} Ka \right|. \end{aligned}$$

Thus the solutions Eq. (3) and the frequencies Eq. (4) corresponding to K and K_m are identical. It means that the state of vibration of the lattice corresponding to a wave vector K is the same as that for any of the wave vectors $K + 2\pi m/a$. Therefore, to find a unique relation between the state of vibration and the wave vector, the wave vector K must be restricted to a range of values $2\pi/a$. We specify the range of independent values of K by

$$-\frac{\pi}{a} \le K \le \frac{\pi}{a}.$$

We have taken both positive and negative values of K because waves can propagate to the right or to the left. This range of K values is referred to as the "first Brillouin zone" of the linear lattice. The extreme values of K in this zone are

$$K_{\max} = \pm \frac{\pi}{a}$$

where K_{max} may be the order of $10^{10}\ m^{-1}$.

At the boundaries $K_{max} = \pm \frac{\pi}{a}$ of the Brillouin zone the solution $x_n = e^{-i(\omega t - Kna)}$ becomes

$$x_n = e^{-i\omega t}\, e^{\pm in\pi} = e^{-i\omega t}(-1)^x \qquad \ldots(8)$$

Eq. (8) represents a "standing" wave in which alternate atoms vibrate in opposite phases. This situation is equivalent to Bragg reflection of X-rays. When the Bragg condition is satisfied, a wave cannot propagate in a lattice, but through successive reflections back and forth, a standing wave is set up. Here the critical value $K_{max} = \pm\frac{\pi}{a}$ satisfies the Bragg condition $2\,d \sin\theta = n\,\lambda$. Here we have $\theta = \frac{\pi}{2}$ (linear lattice), $d = a$, $n = 1$, so that

$$\lambda = \lambda_{min} = 2a$$

which corresponds to $K_{max} = \pm\,\pi / a$.

3. The frequency corresponding to zone boundaries ($K = K_{max} = \pm\,\pi / a$) is according to Eq. (4),

$$\omega_{max} = 2\sqrt{\frac{f}{m}}$$

or
$$\nu_{max} = \frac{1}{\pi}\sqrt{\frac{f}{m}}. \qquad [\because\ \omega = 2\pi\nu]$$

Thus there exists a maximum frequency ν_{max} which can be propagated through the chain. Hence a monatomic linear lattice acts as a low-pass filter which transmits only in the frequency range between zero and ν_{max}.

$$\nu_{max} = \frac{v}{\lambda_{min}} = \frac{v}{2a}.$$

Now, $a = 10^{-10}\,m$. The velocity of sound in solids is $\sim 10^3 - 10^4\ \text{ms}^{-1}$.

$$\therefore \qquad \nu_{max} \approx 10^{13}\ \text{Hz}.$$

Hence the cut off frequency lies in the infrared region.

The actual physical character of motion is illustrated by Fig. 50.3 (*a*) for the long wavelength case. Fig. 50.3 (*b*) illustrates the condition $\lambda = 2a$ in which the stationary waves are set up.

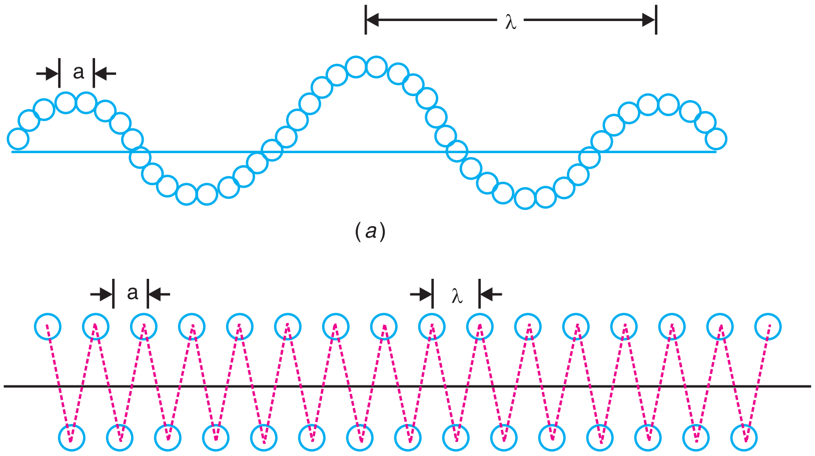


Fig. 50.3

The displacements are shown as transverse ones for the sake of cleârity in representation, but it should be remembered that the actual desplacements which are discussed in the text are longitudinal.

50.3 Vibrational Modes of a Finite One-dimensional Lattice of Identical Atoms

Let there be $N + 1$ atoms in the array numbered from 0 to N. We further suppose that the two end atoms are rigidly fixed so that the number of mobile atoms is $(N - 1)$ (Fig. 50.4). If the length of the chain be L, and each atom separated from its neighbour by a distance a, then $a = L/N$.

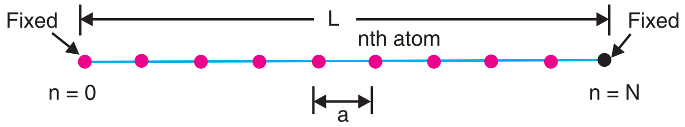


Fig. 50.4

Let a be the equilibrium interatomic spacing, and x_0, x_1, x_2, ... x_N the instantaneous displacements from the equilibrium positions. Let m be the mass of an atom and f the force constant for the nearest-neighbour interaction.

Assuming Hooke's law to be obeyed and considering only nearest-neighbour interaction, the equation of motion of the nth atom is

$$m\ddot{x}_n = f(x_{n+1} - x_n) - f(x_n - x_{n-1})$$

$$= f(x_{n+1} + x_{n-1} - 2x_n) \quad \text{...(1)}$$

The general solution of Eq. (1) can be written as a sum of two travelling waves, one propagating to the right and the other to the left :

$$x_n = \left[A_1 e^{i(Kna+\beta_1)} + A_2 e^{i(-Kna+\beta_2)}\right] e^{-i\omega t} \quad \text{...(2)}$$

where A_1 and A_2 are amplitudes, β_1 and β_2 are phase angles and K is wave vector ($K = 2\pi/\lambda$). The boundary conditions are

$$\left.\begin{aligned} x_0 &= 0 \\ x_N &= 0 \end{aligned}\right\} \text{ for all } t$$

The first boundary condition, when substituted in Eq. (2) yeilds

$$\left.\begin{aligned} A_1 &= -A_2 \\ \text{and} \quad \beta_1 &= \beta_2 = 0, \end{aligned}\right\} \quad \text{...(3)}$$

(since phase angles are equal we choose them to be equal to zero)

Eq. (2) becomes

$$x_n = A_1 (e^{iKna} - e^{-iKna}) e^{-i\omega t} \quad \text{...(4)}$$

Taking real part of Eq. (4), we have

$$x_n = 2A_1 \sin Kna \sin \omega t. \quad \text{...(5)}$$

This represents a "standing" wave and leads to the dispersion relation

$$\omega^2 = \frac{4f}{m} \sin^2 \frac{Ka}{2} \quad \text{...(6)}$$

as that for travelling wave solution [Eq. (4) of Art. 50.2], with the difference that K is now limited to positive values ranging from 0 to π/a.

The second boundary condition imposed on Eq. (5) gives

$$0 = 2A_1 \sin KNa \sin \omega t$$

$$\text{or} \quad \sin KNa = 0$$

$$\text{or} \quad K = \frac{\pi}{Na} j, \quad \text{...(7)}$$

where j is an integer. However, $j = 0$ must be excluded because it corresponds to $K = 0$ which means that all atoms are at rest. The maximum value of K is $\frac{\pi}{a}$ which gives $j_{max} = N$. This value must also be excluded because it yields in Eq. (5) $x_n = 0$ for all t, that is again all atoms are at rest. We, therefore, conclude that

$$j = 1, 2, 3, \ldots\ldots\ldots\ldots, (N-1) \qquad \ldots(8)$$

Thus the boundary conditions select a discrete set of K values given by Eqs. (7) and (8). We see that, *there are just as many modes of vibration* (K-*values*) *as there are mobile atoms.*

To each value of K there corresponds a frequency ω_k. Hence the frequency spectrum consists of $(N-1)$ discrete lines.

50.4 Vibrational Modes of a Diatomic Linear Lattice

For crystals with more than one atom per primitive cell, the vibrational spectrum shows new features. For each polarization mode in a given direction of propagation, the dispersion relation $\omega - K$ gives rise to two branches, known as 'acoustical branch' and 'optical branch'.

Consider a linear diatomic lattice, with atoms of two kinds arranged alternately (Fig. 50.5). The particles are numbered in such a way that the even numbered have mass M and odd ones m $(M > m)$.

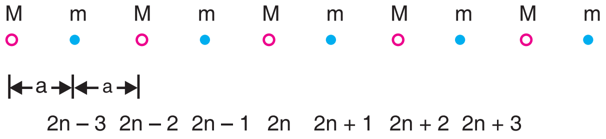


Fig. 50.5

Let a be the distance between nearest neighbours, so that $2a$ is the repeat distance.

We assume that each atom interacts only with its nearest neighbour and that the force constants are identical between all nearest-neighbour pairs. The equations of motion are

$$\left.\begin{aligned} M\ddot{x}_{2n} &= f\,(x_{2n+1} + x_{2n-1} - 2x_{2n}) \\ \text{and} \qquad m\ddot{x}_{2n+1} &= f\,(x_{2n+2} + x_{2n} - 2x_{2n+1}) \end{aligned}\right\} \qquad \ldots(1)$$

where f is the force constant and x_{2n} is the displacement of the $2n$th particle from the equilibrium position.

Let the solutions to these equations be of the form of a travelling wave :

$$\left.\begin{aligned} x_{2n} &= A\,e^{-i(\omega t - 2nKa)} \\ \text{and} \qquad x_{2n+1} &= B\,e^{-i[\omega t - (2n+1)Ka]} \end{aligned}\right\} \qquad \ldots(2)$$

where K is the wave vector of a particular mode of vibration, and A and B are the amplitudes corresponding to the atoms M and m respectively.

Substituting Eq. (2) in Eq. (1), we get

$$\left.\begin{aligned} (M\omega^2 - 2f)\,A + 2Bf\cos Ka &= 0 \\ \text{and} \qquad (m\omega^2 - 2f)\,B + 2Af\cos Ka &= 0 \end{aligned}\right\} \qquad \ldots(3)$$

These equations have non-vanishing solutions for A and B only if the determinant of the coefficients of A and B vanishes, *i.e.*,

$$\begin{vmatrix} (M\omega^2 - 2f) & 2f\cos Ka \\ 2f\cos Ka & (m\omega^2 - 2f) \end{vmatrix} = 0$$

or $$(M\omega^2 - 2f)(m\omega^2 - 2f) - 4f^2\cos^2 Ka = 0$$

or $$Mm\,\omega^4 - 2f(M+m)\,\omega^2 + 4f^2(1 - \cos^2 Ka) = 0$$

or $$Mm\,\omega^4 - 2f(M+m)\,\omega^2 + 4f^2\sin^2 Ka = 0$$

This gives the dispersion relation

$$\omega^2 = f\left(\frac{1}{M}+\frac{1}{m}\right) \pm f\sqrt{\left[\left(\frac{1}{M}+\frac{1}{m}\right)^2 - \frac{4\sin^2 Ka}{Mm}\right]}$$

This is the dispersion relation for the linear diatomic lattice. The form of the relation shows that, as in the monatomic case, the frequency ω is a periodic function of the wave vector K.

Since ω should be positive, each value of ω^2 leads to a single value for ω. Thus, in contrast to the monatomic lattice, there are now two angular frequencies ω_+ and ω_- corresponding to a single value of the wave vector K. In a $\omega - K$ plot (Fig. 50.6) this leads to two "branches", the one corresponding to ω_+ is called the 'optical' branch and that corresponding to ω_- is called the 'acoustical' branch.

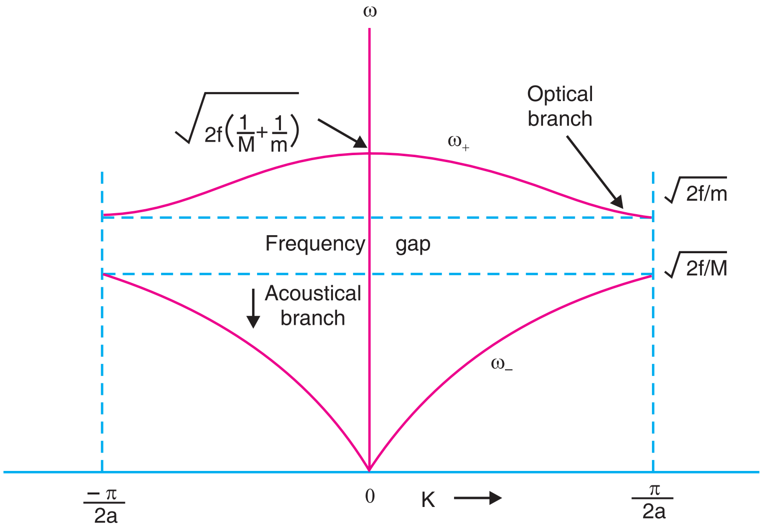


Fig. 50.6

It is further observed that as in the mono-atomic case, the frequency is a periodic function of the wave vector K. The first zone thus limits the values of K to the range between $-\pi/2a$ and $+\pi/2a$ as shown in Fig. 50.6. For $K = \pm\pi/2a$, the frequencies are

$$\omega_+ = \sqrt{\frac{2f}{m}}, \quad \omega_- = \sqrt{\frac{2f}{M}} \qquad \text{...(5)}$$

It is of interest to investigate the physical difference between the acoustical and optical braches and why they are named so. To find out, we calculate $\frac{A}{B}$ as $K \to 0$ for both the branches.

As $K \to 0$, $\cos Ka \to 1$ and $\sin Ka \to 0$

$$\omega_- = 0 \textit{ and } \omega_+ = \sqrt{2f\left(\frac{1}{M}+\frac{1}{m}\right)} \qquad \text{...(6)}$$

$$\left.\begin{aligned} &\text{For acoustical branch} & A &= B \\ &\text{and for optical branch} & -MA &= mB \end{aligned}\right\} \qquad \text{...(7)}$$

For the optical branch at $K = 0$, the vibrations of the atoms are in opposite directions and amplitudes are inversely in the ratio of the masses, so that the centre of mass of the unit cell remains fixed during the peiod of motion. For acoustical mode of the vibration, the two types of atoms move in the same direction with same amplitude as shown in Fig. 50.7. Both waves are travelling waves. The optical branch can have $K = 0$ mode but the acoustical cannot.

These characteristics are typical of the optical and acoustical branch in general. The acoustical branch vibrations can be excited by some kind of force that forces all the atoms in the crystal to go in the same direction. For example, we may direct a beam of sound waves at the surface of a crystal and this will produce the desired effect. It is for this reason, such vibrations are called the acoustical vibrations and constitute the acoustical branch in dispersion relation.

The optical mode of vibrations in ionic crystals where the two types of atoms are oppositely charged, can be excited by an electric field, which tends to move the ions in opposite directions. Specially in ionic crystals, this mode can excited by the electric field associated with light wave, from which the term optical vibrations is derived.

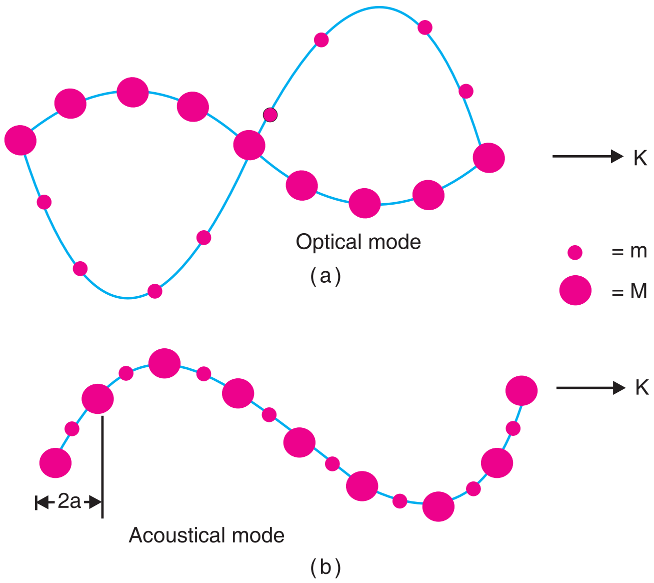


Fig. 50.7

50.5 Quantisation of Lattice Vibrations

The energy in a lattice vibration or elastic wave is quantized. The quantum of energy in an elastic wave is called *phonon*. In fact, phonon represents the particle concept associated with the *quantization of lattice waves.* The situation is quite analogous to the electromagnetic waves where the quanta are called *photons.* It is quite possible to treat the interaction between two phonons or between a phonon and an electron as scattering 'collision' between two particles. Sound waves in crystals are composed of phonons. Thermal vibrations in crystals are thermally excited phonons.

It is convenient to call $\hbar\,\omega_k$ (ω_k is the angular frequency of the kth mode of vibration) the energy of a 'phonon'. The vibrational mode at the temperature T corresponds to

$$n = \frac{1}{e^{\hbar\,\omega/kT} - 1} \qquad \text{...(1)}$$

phonons. At any temperature T the crystal may be regarded as filled with a gas of phonons. More phonons are created with increase in temperature and their number is decreased by lowering temperature. As the phonons are associated with elastic waves in solid, they travel with the velocity of sound (C_s) in the solid medium. The energy of a phonon is

$$\hbar\omega_k = \hbar C_s\, k \qquad \text{...(2)}$$

The vibration spectrum of the phonon waves occupy a wide range of frequency from 10^4 to 10^{13} Hz. The low frequency part of the spectrum is in the acoustic range and the higher frequency part is in infrared range.

The quantisation of lattice vibration energy has an important consequence when lattice vibration is involved in an interaction. The energy of a mode of vibration has to be $(n + 1/2)\,\hbar\,\omega$ with n any positive integer. Further, to a first order approximation, transition occurs only between adjacent allowed states of the system. Therefore, the energy change $\Delta\,\varepsilon$ of the oscillator corresponding to the energy of the quantum which is absorbed or emitted must be

$$\Delta\,\varepsilon = \pm\,\hbar\omega$$

i.e., in any such transition the quantum number of the oscillator may change only by

$$\Delta\,n = \pm 1.$$

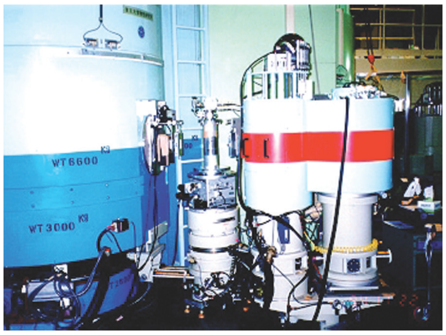
The (+) sign refers to absorption and the (–) sign to emission.

There is no direct experimental evidence that the energy of an elastic wave is quantised. However, the following facts provide evidence for the existence of phonons :

(*i*) The lattice contribution to the heat capacity of solids always approaches zero as the temperature approaches zero. This can be explained only if the lattice vibrations are quantised.

(*ii*) X-rays and neutrons are scattered inelastically by crystals. The change of energy and momentum in this process corresponds to the creation or absorption of one or more phonons.

50.6 Inelastic Scattering of Neutrons by Phonons



Neutron Inelastic Scattering.

A phonon of wave vector K interacts with other particles such as photons, neutrons and electrons as if it has a momentum $\hbar\,K$. Sometimes $\hbar\,K$ is called the 'crystal momentum'.

Neutron inelastic scattering is the most common method for the experimental determination of phonon dispersion relations. A neutron sees the crystal lattice chiefly by interaction with the nuclei of the atoms. The kinematics of the scattering of a neutron beam by a crystal lattice are described by the general wave vector selection rule :

$$\mathbf{k} + \mathbf{G} = \mathbf{k'} \pm \mathbf{K} \qquad \text{...(1)}$$

and by the requirement of the conservation of energy. Here **K** is the wave vector of the phonon created (+) or absorbed (–) in the process, and **G** is any reciprocal lattice vector. We choose **G** such that **K** lies in the first Brillouin zone, as it must for a phonon.

The kinetic energy of the incident neutron is

$$\text{K·E.} = \frac{1}{2}M_n v^2 = \frac{1}{2}\frac{p^2}{M_n} = \frac{\hbar^2 k^2}{2M_n}$$

where M_n is the mass of the neutron. The momentum p is given by $\hbar\mathbf{k}$ where $\mathbf{k}$ is the wave vector of the neutron. If $\mathbf{k}'$ is the wave vector of the scattered neutron, the energy of the scattered neutron is $\frac{\hbar^2 k'^2}{2M_n}$. Thus from the conservation of energy law, we have

$$\frac{\hbar^2 k^2}{2M_n} = \frac{\hbar^2 k'^2}{2M_n} \pm \hbar\omega \qquad \text{...(2)}$$

where $\hbar\,\omega$ is the energy of the phonon created (+) or absorbed (–) in the process.

To determine the dispersion relation using Eqs. (1) and (2), it is necessary in the experiment to find the energy gain or loss of the scattered neutron as a function of the scattering direction $\mathbf{k} - \mathbf{k}'$. We determine **G** the reciprocal lattice vector, from the elastically scattered (Bragg reflected) neutrons. The elastic scattering of a neutron by crystals obeys the following wave vector conservation law :

$$\mathbf{k}' = \mathbf{k} + \mathbf{G} \qquad \text{...(3)}$$

where **k**′ is wave vector of scattered neutron, and **k** wave vector of incident neutron.

The reciprocal lattice vector **G** can be measured through Eq. (3) by elastic scattering for a given **k** and **k**′. This, in turn, gives **K** in Eq. (1). On the other hand, ω is obtained from Eq. (2). Thus a pair (**K**, ω) could be obtained. This would locate a point on the dispersion curve. By changing **k** (and hence **k**′), other points on the dispersion curve could be located to get the entire graph.

As an example of a widely used experimental method for the measurement of phonon dispersion relation, we briefly describe the triple axes (neutron) spectrometer (TAS), shown schematically in Fig. 50.8. The wave vector (**k**) of the incoming neutron beam is known from the reactor characteristics. The TAS method is designed to measure the magnitude and the direction of the scattered beam (k') through a facility for changing the angle (ϕ) between **k** and **k**′ and the angle (ψ) between k and some reference direction of the crystal.

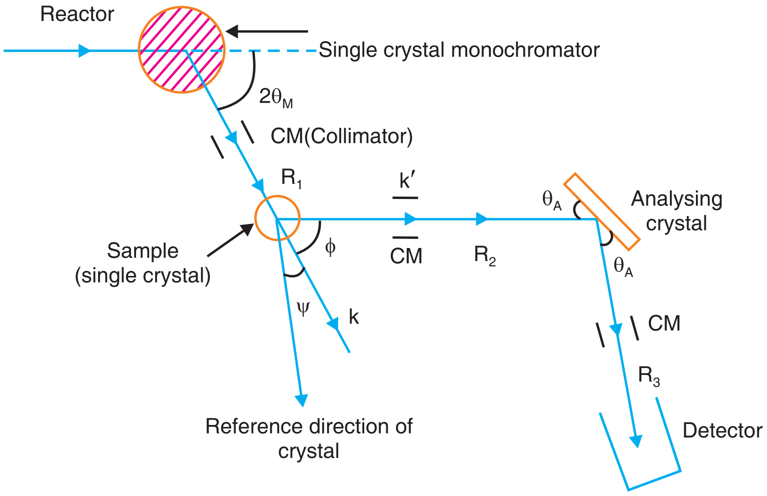


Fig. 50.8

By changing the angle θ_M, it is possible to change **k**. By a proper setting of the table of the crystal (angle ψ), any desired angle between **k** and a certain crystal plane (characterised by a set of Miller indices) corresponding to some reciprocal lattice vector may be obtained. By rotating the second arm, the angle ϕ between **k** and **k**′ can be changed. By rotating the third arm (with the crystal analyser synchronised to follow with half the angular velocity) the scattered beam in the direction **k**′ can be analysed.

Triple Axes Spectrometer.

The neutrons of desired energy are scattered towards the sample crystal by Bragg reflecting them from a suitably oriented single crystal (the monochromator in Fig. 50.8). The single crystal analyser is used to define the energy of those neutrons scattered from the sample and which reach the detector. The whole arrangement is called the triple axes spectrometer, since rotation about three axes is required to adjust the three scattering angles.

The dispersion curves of sodium for phonons propagating in [100], [110] and [111] directions at 90 K, as determined by inelastic scattering of neutrons are given in Fig. 50.9.

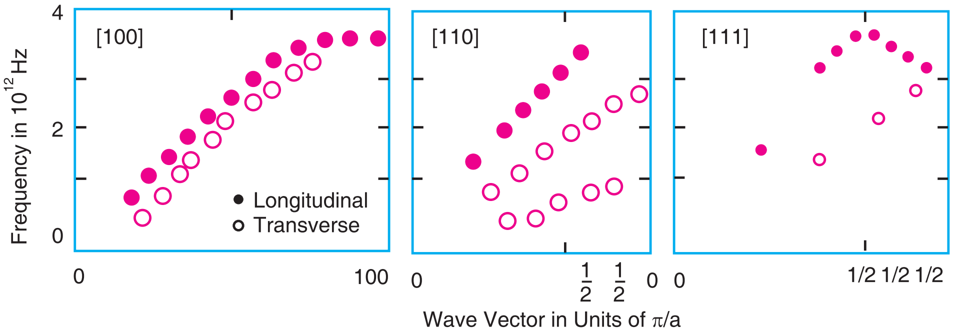


Fig. 50.9

EXERCISE

1. Discuss the vibrations of monatomic linear lattices and compare the results with the vibrations of a continuous string.
2. Discuss the modes of vibration of a finite one-dimensional lattice of identical atoms.
3. Discuss the vibrations of a diatomic linear lattice. Show that the vibrational spectrum consists of two branches. What are the main features of these branches?
4. Derive the vibrational modes of a diatomic linear lattice. What is the difference between the two branches and why are they named so? [*Indore,* 1980)
5. Discuss the vibrations of diatomic lattice and describe its optical and acoustic modes. [*Delhi,* 1983]
6. What is phonon? How are the conservations of energy and momentum expressed in the case of inelastic scattering of a phonon by (*i*) a photon, and (*ii*) a neutron? Explain why the phonon energy can be measured much more accurately by the method of scattering of neutrons than by that of X-rays.
7. Explain how one obtains experimentally information about the phonon dispersion curves by using the neutron scattering technique. Inter compare this method with the method of X-ray scattering and infrared absorption in certain systems. [*Udaipur,* 1982]
8. Sound velocities in solids are of the order of 3×10^3 ms^{-1}. Interatomic distances in solids are of the order of 3×10^{-10} m. Estimate the order of magnitude of cut-off frequency, assuming a linear lattice.

[**Ans.** 3×10^{12} Hz]

CHAPTER 51

ELECTRICAL CONDUCTIVITY AND OHM'S LAW

AT A GLANCE

51.1 Electrical Conductivity
51.2 Experimental Electrical Resistivity of Metals
51.3 Conduction Electrons
51.4 Electron Collision
51.5 Derivation of Boltzmann Transport Equation
51.6 Paramagnetic Susceptibility of Conduction Electrons

51.1 Electrical Conductivity

The momentum of a free electron is related to the wavevector by

$$mv = \frac{l}{\hbar}k \qquad \text{...(1)}$$

In an electric field **E** and magnetic field **B**, the force **F** on an electron of charge – *e* is

$$-e[\mathbf{E}+\mathrm{v}\times\mathbf{B}] \qquad \text{...(2)}$$

Newton's second law of motion becomes

$$\boldsymbol{F} = m\frac{d\boldsymbol{v}}{dt} = \hbar\frac{d\mathbf{k}}{dt} = -e(\mathbf{E}+\boldsymbol{v}\times\mathbf{B})$$

Let $\mathbf{B} = 0.$

$$\hbar\frac{d\mathbf{k}}{dt} = -e\mathbf{E} \qquad \text{...(3)}$$

Integrating, $$\mathbf{k}(t)-\mathbf{k}(0) = -\frac{e\mathbf{E}t}{\hbar} \qquad \text{...(4)}$$

In the absence of collisions the Fermi sphere (Fig. 51.1) in **k** space is displaced at a uniform rate by a constant applied electric field.

Let us consider electrical conductivity in **k** space.

***(i)* In the absence of an electric field.** In the absence of an electric field, the Fermi sphere is centered at the origin (Fig. 51.1 *a*).

The Fermi sphere encloses the occupied electron orbitals in **k** space in the ground state of the electron gas. The net momentum is zero, because for every orbital **k** there is an occupied orbital at – **k**.

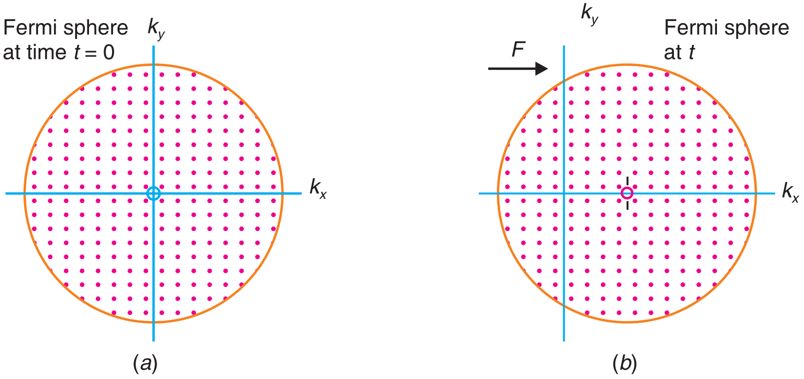
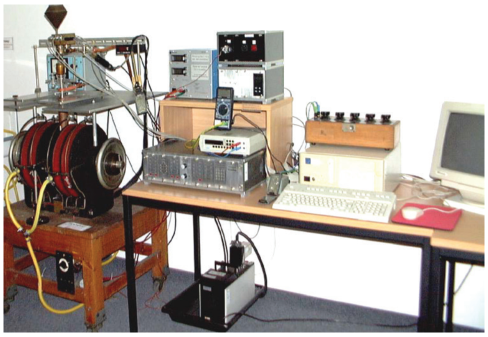


Fig. 51.1

***(ii)* When an electric field is applied.** Suppose the electric field is applied at time $t = 0$ to an electron gas that fills the Fermi sphere centered at the origin of **k** space. Then, at a later time t the sphere will **be** displaced to a new centre at

$$\delta \mathbf{k} = -\mathbf{E}t/\hbar \qquad ...(5)$$



Determination of Electrical Conductivity.

Notice that the Fermi sphere is displaced as a whole. Thus in **k** space electrical conduction is associated with a slight displacement of the Fermi sphere about the origin.

Because of collisions of electrons with impurities, lattice imperfections and phonons, the displaced sphere may be maintained in a steady state in an electric field.

Let τ = collision time.

Then, the displacement of the Fermi sphere in the steady state is given by Eq. (5) with $t = \tau$.

The incremental velocity is

$$v = -eE\tau/m \qquad ...(6)$$

Let there be n electrons of charge $q = -e$ per unit volume.

The electric current density is

$$\boldsymbol{j} = \mathrm{nqv} = ne^2\tau\mathbf{E}/m \qquad ...(7)$$

This is Ohm's law.

The electrical conductivity σ is defined by $\boldsymbol{j} = \sigma \boldsymbol{E}$.

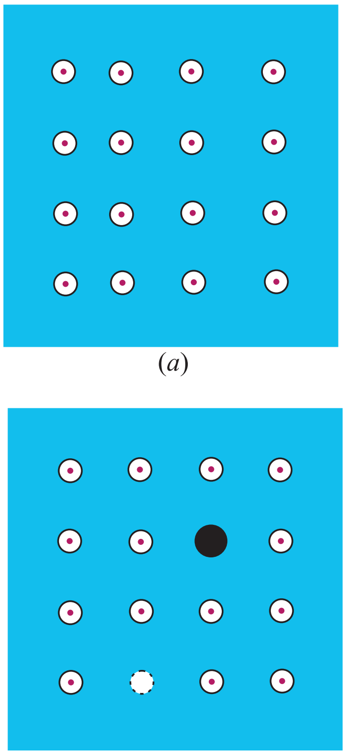
$$\therefore \qquad \sigma = \frac{ne^2\tau}{m} \qquad ...(8)$$

The electrical resistivity ρ is defined as the reciprocal of the conductivity.

$$\therefore \qquad \rho = \frac{m}{ne^2\tau} \qquad ...(9)$$

It is easy to understand the result Eq. (8) for the coductivity. We expect the charge transported to be proportional to the charge density *ne*. The factor *e*/*m* enters because the acceleration in a given electric field is proportional to *e* and inversely proportional to the mass *m*. The time τ describes the free time during which the field acts on the carrier.

51.2 Experimental Electrical Resistivity of Metals


(*a*)

(*b*)

Fig. 51.2

The electron undergoes a collision only because the lattice is not perfectly regular. It is convenient to group the deviations from a perfect lattice into two classes :

(a) Lattice vibrations (phonons) of the ions around their equilibrium position due to thermal excitations of the ions [Fig. 51.2(a)].

(b) All static imperfections, such as foreign impurities and vacant lattice sites [Fig. 51.2(*b*)].

The electrical resistivity of most metals is dominated at room temperature (300 K) by collisions of the conduction electrons with lattice phonons and at liquid helium temperature (4K) by collisions with impurity atoms and mechanical imperfections in the lattice.

To find the net resistivity ρ :

(*i*) The probabilities of electrons being scattered by phonons and by impurities are additive, since these two mechanisms are assumed to act independently.

(*ii*) Since the scattering probability is inversely proportional to the relaxation time, the reciprocal on the effective relaxation time is the sum of the reciprocals of the individual relaxation times.

Therefore, we write

$$\frac{1}{\tau} = \frac{1}{\tau_L} + \frac{1}{\tau_i} \qquad \text{...(1)}$$

Here, the first term on the right is due to phonons and the second is due to impurities.

(*i*) τ_L is expected to depend on *T*.

(*ii*) τ_i depends on impurities, but not on *T*.

The general formula for electrical resistivity is

$$\rho = 1/\sigma = m/ne^2\tau \qquad \text{...(2)}$$

The net resistivity is given by

$$\rho = \rho_L + \rho_i = \frac{m}{ne^2}\frac{1}{\tau_L} + \frac{m}{ne^2}\frac{1}{\tau_i} \qquad \text{...(3)}$$

Electrical Resistivity Test Equipment.

(*i*) ρ_L is the resistivity caused by the thermal phonons. Hence it is temperature dependent, and is called the *ideal resistivity*. It is the resistivity of a pure specimen.

The temperature-dependent part of the electrical resistivity is proportional to the rate at which an electron collides with thermal phonons and thermal electrons.

Often ρ_L is independent of the number of defects when their concentration is small.

(*ii*) ρ_i is the resistivity caused by scattering of the electron waves by static defects that disturb the periodicity of the lattice. Often ρ_i is independent of temperature. The residual resistivity, ρ_i (0), is the extrapolated resistivity at 0K because ρ_L vanishes as $T \rightarrow 0$.

Fig. 51.3 shows resistance of potassium below 20K, as measured on two specimens. The different intercepts at 0K are attributed to different concentrations of impurities and static imperfections in the two specimens.

The lattice resistivity, $\rho_L(T) = \rho - \rho(0)$, is the same for different specimens of a metal, even though ρ_i (0) may itself vary widely. *The resistivity ratio* of a specimen is defined as the ratio of its resistivity at room temperature to its residual resistivity. It is a convenient approximate indicator of sample purity. In exceptionally pure specimens the resistivity ratio may be as high as 106, whereas in some alloys (*e.g.*, manganin) it may be as low as 1.1.

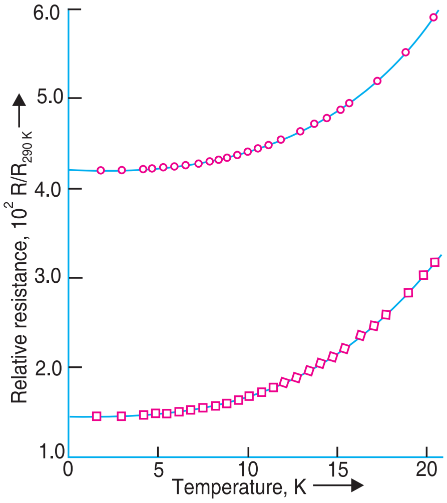
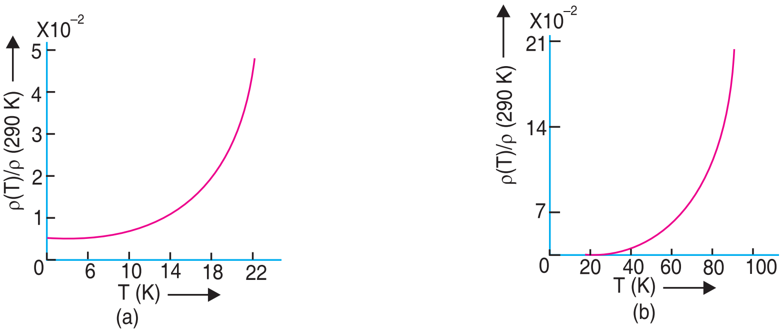


Fig. 51.3

Matthiessen's rule. Eq. (3) tells us that the resistivity in the presence of several distinct scattering mechanisms is simply the sum of the resistivities one would have if each alone were present. This is called *Matthiessen's rule*. In other words, this rule states that ρ can be split into two parts, one of which is independent of *T*. Matthiessen's rule is convenient in analyzing experimental data (Fig. 51.3).

Temperature Dependence of Electrical Resistivity of Metals

In a pure metal, we except that $\rho \rightarrow 0$ as $T \rightarrow 0$. However, experimental observations show that ρ for metals decreases to a small but finite value and remains practically constant and independent of temperature in the range 10 *K* to 0*K* (Fig.51.4).



(a)

(b)

Fig. 51.4

The resistivity at 0K is called the *residual resistivity* of the metal. This resistivity may be attributed to

(*i*) presence of impurity atoms in small concentrations, and

(*ii*) geometrical imperfections like point defects, grain boundaries and stacking faults in the metal.

Fig. 51.4 shows the normalised resistivity ρ (*T*)/ρ(290 K) versus *T* for Na (*a*) in low temperature region and (*b*) at higher temperatures, $(290) \simeq 2.10 \times 10^{-8}\ \Omega\text{-}m$.

(*a*) At very low temperatures, scattering by phonons is negligible, because the amplitudes of oscillations are very small in that region.

$$\tau_L \to \infty, \rho_L \to 0.$$

$$\therefore \qquad \rho = \rho_i = \text{a constant.}$$

This is in agreement with Fig. 51.4 (*a*).

(*b*) As T increases, scattering by phonons becomes more effective, and $\rho_L(T)$ increases ; this is why ρ increases. When T becomes sufficiently large scattering by phonos dominates and $\rho = \rho_L$ (T). In the high temperature region, $\rho_L(T)$ increases linearly with T. This is again in agreement with experimental results shown in Fig. 51.4 (b).

We expect that ρ_i should increase with impurity concentrations, and indeed it could be shown that for small concentrations ρ_i is proportional to the impurity concentration N_i. We also remark that, for small impurity concentration $\rho_L >> \rho_i$ except at very low temperature.

MOTION OF ELECTRONS

51.3 Conduction Electrons

A metallic solid contains a large number of electrons which are free to move throughout the solid. These electrons give rise to the electrical conductivity of a metal. These electrons are called *conduction electrons*.

(*i*) When free atoms form a metal, all the valence electrons become conduction electrons and their states are profoundly modified.

(*ii*) The core electrons remain localised and their character remains essentially unchanged.

(*iii*) Conduction electrons are responsible for most of the properties of metals.

(*iv*) The electric current in the metal is due to the motion of the conduction electrons under the influence of the electric field. We say that it is the conduction electrons which are responsible for the current because the positive ions are attached to and vibrate about the lattice sites. They have no net translational motion, and hence do not contribute to the current.

(*v*) We can calculate the number of conduction electrons from the valence of the metal and its density.

Electron concentration, $$n = z\frac{\rho_m N}{A}.$$

Here,
z = atomic valence,
ρ_m = density of the substance,
N = Avogadro's number, and
A = atomic weight.

51.4 Electron Collisions

Conduction electrons move throughout the metal. They continually interact with the atoms as well as with each other. We treat the conduction electrons as point masses which are moving about in a "box" filled with "obstacles" of some kind. The electrons frequently "collide" with these obstacles. Thus, the electrons experience collisions (by an unspecified interaction), or that the electrons are scattered.

(*i*) The collisions are treated as instantaneous scattering events. The time for the scattering to take place is much shorter than any other times in the problem. It is through these collisions that electrons achieve thermal equilibrium corresponding to the metal temperature T.

(*ii*) **Relaxation time, Collision time and Mean free path**

(*a*) The time between two successive collisions in the presence of applied electric field is defined as *relaxation time* (τ) of the electrons.

(*b*) *Collision time* τ_c is the average time between two successive collisions.

For an isotropic solid, like a metal $\tau = \tau_c$.

(*c*) We define the mean free path as the average distance travelled by an electron between collisions. It is given by

$$l = v_m \tau.$$

Here, v_m = root mean square velocity.

The mean free path l of a conduction is defined as

$$l = v_F \tau.$$

Here, v_F is the velocity at the Fermi surface, because all collisions involve only electrons near the Fermi surface.

(*iii*) After each collision, the electron has no memory of what went before. The electrons emerge from collisions with no memory of their velocities before and are then randomly directed with velocities appropriate to T.

(*iv*) ***Drift Velocity***

Between collisions the electrons travel in straight lines obeying Newton's laws.

If an electric field $\vec{E}$ is applied in the x-direction, then we have

$$m\ddot{x} = -e\vec{E} \qquad ...(1)$$

Thus the electron will have an additional velocity given by

$$-\left(\frac{e\vec{E}}{m}\right)t$$

for as long as the electric field is applied and as long as the electron is not scattered.

However, on the average, the electrons are scattered after a time τ. After each scattering event they are in thermal equilibrium. So a constant electric field will cause the electrons to have an extra average velocity given by

$$v_d = -\left(\frac{e\vec{E}}{m}\right)\tau \qquad ...(2)$$

This is the drift velocity, v_d, that is due to the applied electric field.

Definition. *The drift velocity "v_d" is the average velocity acquired by the free electron of a metal in a particular direction during the application of the electric field.*

The root mean square velocity, v_m, in most metals is very much larger than v_d, *i.e.*,

$$v_m >> v_d.$$

51.5 Derivation of Boltzmann Transport Equation

The classical theory of transport processes is based on the Boltzmann transport equation. Consider a region of phase space about the point (x, y, z, p_x, p_y, p_z). The number of particles entering this region in time dt is equal to the number which were in the region of phase space at

$$\left(x - v_x dt, y - v_y dt, z - v_z dt,\ z - v_z dt,\ p_x - F_x dt, p_y - F_y\, dt, p_z - F_z dt\right)$$

at a time dt earlier.

Ludwig Boltzmaun
(1844-1906)

Here $\boldsymbol{F}$ represents the force acting on the particles of the distribution at the point (x, y, z) and time t.

From Newton's law,

$$F = \frac{dp}{dt}.$$

Let $f(x, y, z, p_x, p_y, p_z)$ be the distribution function which expresses the number of particles per quantum state in the region. Then the change df which occurs during time dt due to the motion of the particles in coordinate space and due to the fact that force fields acting on the particles tend to move them from one region to another in momentum space is

$$df = f(x - v_x dt, y - v_y\, dt,\ z - v_z\, dt, p_x - F_x\, dt, p_y - F_y dt, p_z - F_z dt) - f(x, y, z, p_x, p_y, p_z) \quad ...(1)$$

Using Taylor's expansion and retaining only first order terms in the limit $dt \to 0$, Eq.(1) may be written as

$$df = \left(-v_x \frac{df}{dx} - v_y \frac{df}{dy} - v_z \frac{df}{dz} - F_x \frac{df}{dp_x} - F_y \frac{df}{dp_y} - F_z \frac{df}{dp_z}\right)$$

i.e.,
$$\frac{\partial f}{\partial t} = -\boldsymbol{v}.\ \text{grad}_r f - \mathbf{F}.\ \text{grad}_p f \quad ...(2)$$

This is *Boltzmann's transport equation* when no collisions are taken into account.

Here, $\text{grad}_r = \nabla_r$ is the usual del operator and

$\text{grad}_p = \nabla_p$ is the del operator in momentum space.

$$\left.\begin{aligned} \boldsymbol{\nabla}_r &= \mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y} + \mathbf{k}\frac{\partial}{\partial z} \\ \text{and} \quad \boldsymbol{\nabla}_p &= \mathbf{i}\frac{\partial}{\partial p_x} + \mathbf{j}\frac{\partial}{\partial p_y} + \mathbf{k}\frac{\partial}{\partial p_z} \end{aligned}\right\} \quad ...(3)$$

$$\therefore \quad \left.\begin{aligned} \mathbf{v}.\text{grad}_r\, f &= v_x \frac{\partial}{\partial x} + v_y \frac{\partial}{\partial y} + v_z \frac{\partial}{\partial z} \\ \text{and} \quad \mathbf{F}.\text{grad}_p f &= \mathbf{F}_x \frac{\partial}{\partial p_x} + \mathbf{F}_y \frac{\partial}{\partial p_y} + \mathbf{F}_z \frac{\partial}{\partial p_z} \end{aligned}\right\} \quad ...(4)$$

So far, only the change in the distribution function due to the motion of the particles in coordinate space and due to the momentum changes arising from force fields acting on the particles have been accounted for.

Particles may also be transferred into or out of a given region of phase space by *collisions* or *scattering interactions* involving other particles of the distribution or scattering centers external to the assembly of particles under consideration.

Let the rate of change of the distribution function due to collision or scattering be denoted by

$$\left(\frac{\partial f}{\partial t}\right)_{\text{collision}}.$$

Then, the total rate of change of f is obtained by adding this quantity to the right-hand side of Eq.(2).

$$\therefore \quad \frac{df}{dt} = -\mathbf{v}.\nabla f - \mathbf{F}.\nabla_p f + \left(\frac{\partial f}{\partial t}\right)_{\text{coll}} \quad ...(5)$$

Eq. (5) is *Boltzmann's transport equation* in general form.

The Boltzmann equation is more commonly used to discuss the properties of a system which is displaced from equilibrium by a small perturbing force, such as an electric field or a temperature gradient, when the distribution function f_0 which is the equilibrium distribution function in the absence of the perturbation is known. In such cases, it is frequently assumed that the collision term has the form

$$\left(\frac{\partial f}{\partial t}\right)_{coll} = -\frac{f-f_0}{\tau} \qquad ...(6)$$

Here, τ is relaxation time.

The Boltzmann transport equation in the relaxation time approximation is

$$\frac{df}{dt}+\mathbf{v}.\nabla f+\mathbf{F}.\nabla pf = -\frac{f-f_0}{\tau}. \qquad ...(7)$$

51.6 Paramagnetic Susceptibility of Conduction Electrons

The *magnetic susceptibility* χ of a specimen is defined as magnetic moment per unit magnetising field *i.e.*,

$$\chi = \frac{M}{H} \qquad ...(1)$$

Here, M is magnetic moment and H is magnetising field.

Classical free electron theory gives an unsatisfactory account of the paramagnetic susceptibility of the conduction electrons. Classically, the *paramagnetic susceptibility* is given by

$$\chi_p = \frac{n\mu_B^2}{k_B T} m^{-3} \qquad ...(1)$$

Here, n = number of conduction electrons per unit volume,

μ_B = Bohr magneton, k_B = Boltzmann constant and T = temperature.

This shows that $\chi_p \propto (1/T)$.

It is observed that the paramagnetic susceptibility of most normal nonferromagnetic metals is independent of temperature.

Pauli used Fermi-Dirac statistics to explain this disagreement between theory and experiment.

We first give a qualitative explanation of the situation.

(*i*) In the absence of any external field ($H = 0$) and at absolute zero ($T = 0$), all energy levels below Fermi level $E_F(0)$ are occupied and all those above $E_F(0)$ are empty [Fig. 51.5(a)]. The electrons are divided into two groups with spin up and spin down.

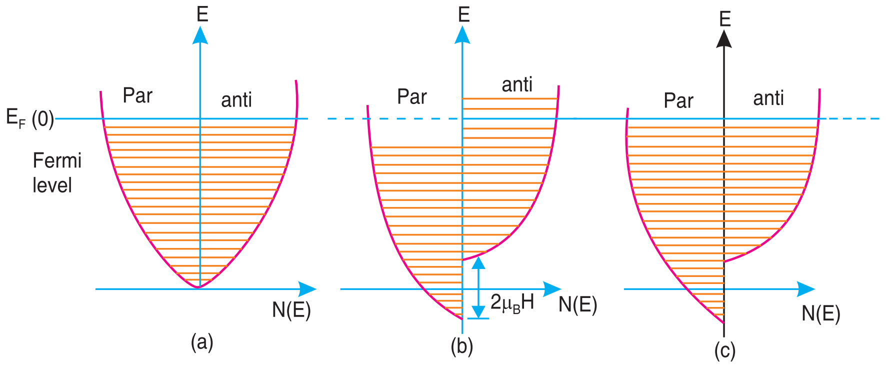


Fig. 51.5

(*ii*) Now, when a magnetic field is applied, the electrons with their spins parallel to the field have their energy lowered by an amount $\mu_B H$ and the electrons with their spins antiparallel raised by the same amount. Fig. 51.5(b) shows instantaneous shift of states when field is applied. It may be noted that $\mu_B H << E_{F0}$. This situation is unstable because for equilibrium the two halves must be filled upto the same level.

(*iii*) Equilibrium is established when a certain number of electrons with antiparallel spins has entered the group of parallel ones till both the halves are filled upto the same level [Fig. 51.5(c)]. The number of such electrons is very small at 0K, and is of the order of $n\left(\frac{T}{T_F}\right)$ per unit volume at temperatures such that $k_B T \ll E_F$, T_F being the Fermi temperature. It is clear that these electrons would lie near the Fermi level. Now, each of these electrons behaves, qualitatively, as if it were to contribute the susceptibility of $\mu_B{}^2 / k_B T$.

The susceptibility of the metal is

$$\chi_p \approx \frac{n\mu_B{}^2}{k_B T_F} \qquad ...(2)$$

This is independent of temperature and is of the observed order of magnitude.

Thus quantum theory is capable of explaining the paramagnetism of metals. We now calculate χ_p quantitatively.

We suppose that the concentration of electrons with their spins, and hence magnetic moments, parallel to the magnetic field is n_p and that of antiparallel ones is n_a. Then expression for n_p may be written as

$$n_p = \int_{-\mu_B H}^{E_F} f(E)\frac{1}{2} D_P (E+\mu_B H)\, dE. \qquad ...(3)$$

Here, $\frac{1}{2}D(E+\mu_B H)$ is the density of states of parallel spin orientation with allowance for the downward shift of energy by $-\mu_B H$.

$f(E) = 1$ throughout the interval of integration at $T = 0$ and

$$D_p(E+\mu_B H) = \frac{1}{2\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2} (E+\mu_B H)^{1/2} \text{ / unit volume}$$

$$\therefore \qquad n_p = \frac{1}{2}\left(\frac{1}{2\pi^2}\right)\left(\frac{2m}{\hbar^2}\right)^{3/2} \int_{-\mu_B H}^{E_F} (E+\mu_B H)^{1/2}\, dE \qquad ...(4)$$

Similarly,

$$n_a = \frac{1}{2}\left(\frac{1}{2\pi^2}\right)\left(\frac{2m}{\hbar^2}\right)^{3/2} \int_{\mu_B H}^{E_F} (E-\mu_B H)^{1/2}\, dE \qquad ...(5)$$

The magnetisation M is given by

$$M = \mu_B(n_p - n_a)$$

$$= \mu_B\left(\frac{1}{2}\right)\left(\frac{1}{2\pi^2}\right)\left(\frac{2m}{\hbar^2}\right)^{3/2}\left[\int_{\mu_B H}^{E_F} (E+\mu_B H)^{1/2}\, dE - \int_{-\mu_B H}^{E_F} (E-\mu_B H)^{1/2}\, dE\right].$$

Since $\mu_B H << E_F$ and there are very few states near $E = 0$, we can set the lower limit in both the integrals equal to zero. Hence

$$M = \left(\frac{\mu_B}{2}\right)\left(\frac{1}{2\pi^2}\right)\left(\frac{2m}{\hbar^2}\right)^{3/2}\left[\int_0^{E_F}(E+\mu_B H)^{1/2}\,dE-\int_0^{E_F}(E-\mu_B H)^{1/2}\,dE\right] \quad ...(6)$$

Now, $$(E+\mu_B H)^{1/2} = \sqrt{E}\left(1+\frac{\mu_B H}{E}\right)^{1/2} = \sqrt{E}+\frac{1}{2}\frac{\mu_B H}{\sqrt{E}}$$

and $$(E+\mu_B H)^{1/2} = \sqrt{E}\left(1-\frac{\mu_B H}{E}\right)^{1/2} = \sqrt{E}-\frac{1}{2}\frac{\mu_B H}{\sqrt{E}}$$

$$(E+\mu_B H)^{1/2}-(E-\mu_B H)^{1/2} = \frac{\mu_B H}{\sqrt{E}}\ .$$

Eq.(6) becomes,

$$M = \left(\frac{\mu_B}{2}\right)\left(\frac{1}{2\pi^2}\right)\left(\frac{2m}{\hbar^2}\right)^{3/2}\mu_B H\int_0^{E_F}\frac{dE}{\sqrt{E}} = \left(\frac{\mu_B^2}{2\pi^2}\right)\left(\frac{2m}{\hbar^2}\right)^{3/2}E_F^{1/2}.H$$

Paramagnetic susceptibility,

$$\chi_p = \frac{M}{H} = \left(\frac{1}{2\pi^2}\right)\left(\frac{2m}{\hbar^2}\right)^{3/2}\mu_B^2\,E_F^{1/2}$$

Substituting the value of E_F, we finally have

$$\chi_p = \frac{3\pi^2 n\mu_B^2}{2\pi^2 E_F} = \frac{3}{2}.\frac{n\mu_B^2}{k_B T_F} \quad ...(7)$$

This is the Pauli paramagnetism formula. This formula is independent of temperature in agreement with experiment. This formula is satisfactory, so long as $k_B T << k_B T_F$.

CHAPTER

52

FREE ELECTRON MODEL OF METALS

AT A GLANCE			
52.1	Sommerfeld's Model	**52.2**	Fermi-Energy
53.3	Sommerfeld Theory of Electrical Conductivity		

52.1 Sommerfeld's Model

Sommerfeld modified the free electron model by taking into account quantum statistics and the Pauli exclusion principle. The assumptions of Sommerfeld model are:

1. The free electrons are the valence electrons of the composing atoms. Thus the alkali metals are assumed to contain one free electron per atom while aluminium has three free electrons per atom.

2. A valence electron in metal finds itself in the field of all ions and that of all other electrons. In this model, the mutual repulsion between the electrons is neglected and the potential field representing the attractive interaction of ions is assumed to be completely uniform everywhere inside the solid.

3. Since the electron emission can not be observed at room temperature, the potential energy of an electron at rest inside the metal is assumed to be lower than that of an electron outside the metal.

Thus the *potential energy for an electron is periodic.* The free electrons find themselves in a potential which is constant everywhere inside the metal but the potential outside the metal is very high which restrains the conduction electrons to remain inside the metal at room temperature. Thus the metal surfaces are the sites of a sharp potential barrier of height Es as shown Fig. 52.1. This figure represents a *physical model for metal* where the interior is represented as consisting of a gas of electrons confined by potential barriers within a certain region of space called a '*potential energy box*' of depth E_s. Thus E_s represents the energy difference between an electron at rest inside the metal and one at rest in vacuum (*i.e.*, on the surface). At $T = 0$, all energy levels upto E_F are filled while all higher ones are empty. Thus work function of metal $\phi = E_S - E_F$.

E_F is called the *Fermi energy and is defined as the maximum kinetic energy that a free electron can have at the absolute zero of temperature.*

4. The 'free electron gas' in a solid does not obey the classical Maxwell Boltzmann statistical law. The distribution of energy in an electron gas obeys the *Fermi-Dirac quantum statistics.*

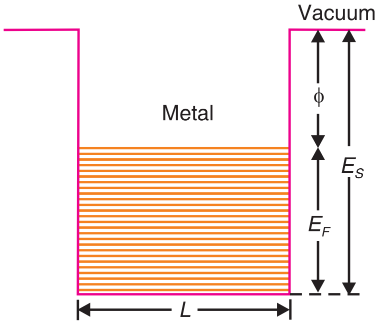


Fig. 52.1

5. In a metal, the energy levels are filled in accordance with the *Pauli's exclusion principle according to which an energy level can accommodate at the most two electrons, one with spin up and the other with spin down. Thus in filling the energy levels, two electrons occupy the lowest level, two more the next level and so forth until all the electrons in the metal have been accommodated.*

Assuming free electrons within a metal ($V_0 = 0$) and using solution of Schrodinger equation for this case, we get the allowed energy values of electrons within the metal as

$$E_n = \frac{n^2 \pi^2 \hbar^2}{2m L^2} \qquad (n = 1, 2, 3...)$$

These are known as *energy eigenvalues*.

The eigenfunction is

$$\psi_n(x) = \sqrt{\left(\frac{2}{L}\right)} \sin \frac{n \pi x}{L}$$

As electron can occupy two spin states there are two quantum states for a single electron. Therefore, density of states for this case would be given by

$$D(E_n) = \frac{dn}{dE_n} = 2.\frac{L}{\pi \hbar}\left(\frac{m}{2 E_n}\right)^{1/2}$$

This result is plotted in Fig. 52.2.

If we consider the free electron gas in three dimensions and extend the above mathematical formulation, we have Schrodinger equation in three dimensions as

$$\nabla^2 \psi + \frac{2m}{\hbar^2} E \psi = 0 \text{ (Since } V_0 = 0)$$

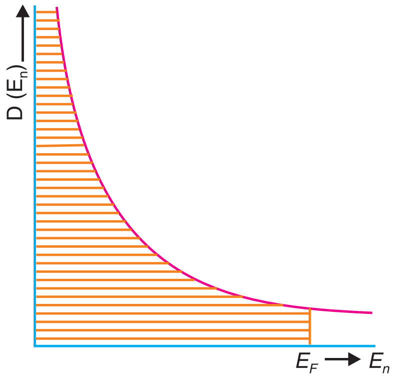


Fig. 52.2

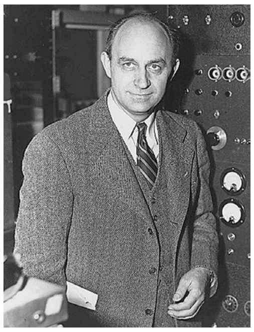
Consider the electrons to be confined in a cube of length L. Then, energy eigenvalues

$$E\,(n_x, n_y, n_z) = \frac{\pi^2 \hbar^2}{2m L^2}(n_x^2 + n_y^2 + n_z^2)$$

$$\text{eigenfunction } \psi\,(x, y, z) = \sqrt{\left(\frac{8}{L^3}\right)} \sin \frac{n_x \pi x}{L} \sin \frac{n_y \pi y}{L} \sin \left(\frac{n_z \pi z}{L}\right)$$

52.2 Fermi-Energy

Let us consider N-non-interacting electrons contained in a box at absolute zero temperature. Using Pauli's exclusion principle on the energy distribution of conduction electrons, it can be shown that all the energy levels below a certain level will be filled with electrons and all the levels above it will be empty. The level which divides the filled and empty levels is called the *Fermi-level at absolute zero* and the corresponding energy of that level is called the *Fermi-energy* and is denoted by $\varepsilon_F(0)$.



Enrico Fermi (1901 – 1954)

Let N non-interacting electrons be distributed among quantum states such that there are $n_1, n_2, \ldots n_i \ldots$ number of particles in quantum states possessing constant energies $\varepsilon_1, \varepsilon_2, \ldots \varepsilon_i \ldots$ respectively. Since electrons are Fermi particles, they obey Pauli's exclusion principle and Fermi-Dirac distribution law.

According to this law, the number of electrons n_i with the energy ε_i is given by

$$n_i = \frac{g_i}{e^{\alpha} e^{\varepsilon_i/kT} + 1} \qquad \ldots(1)$$

where g_i is the number of quantum states having the same energy ε_i.

The number of particle states lying between momentum p and $p + dp$ is given by

$$g(p)\,dp = g_s \cdot \frac{4\pi p^2}{h^3/V} dp = g_s \cdot \frac{4\pi V p^2\,dp}{h^3} \qquad \ldots(2)$$

where g_s is the *spin degeneracy factor* caused by the particle spin s.

$$g_s = (2s + 1) = 2 \text{ for electron gas since } s = 1/2 \qquad \ldots(3)$$

This means that each level can be occupied by two particles one with spin up and the other with spin down (↑↓).

Now $\quad \varepsilon = p^2/2m$ or $dp = (m/2\varepsilon)^{1/2}\,d\varepsilon$.

The number of states between energy ε and $\varepsilon + d\varepsilon$ is given by

$$g(\varepsilon)\,d\varepsilon = 2\frac{4\pi V(2m\varepsilon)\left(\dfrac{m}{2\varepsilon}\right)^{1/2} d\varepsilon}{h^3} \qquad \ldots(4)$$

There is a continuous distribution of electron energies in the solid. In this situation the F-D distribution law becomes

$$n(\varepsilon)\,d\varepsilon = \frac{g(\varepsilon)\,d\varepsilon}{e^{\alpha} e^{\varepsilon/kT} + 1}, \qquad \ldots(5)$$

Here $n(\varepsilon)\,d\varepsilon$ is the number of electrons having energies between ε and $\varepsilon + d\varepsilon$.

For a gas of fermions at temperature T in a volume V, the mean number of fermions in the energy range between ε and $\varepsilon + d\varepsilon$ is given by

$$n(\varepsilon)\,d\varepsilon = \frac{8\pi m V}{h^3}\sqrt{2m} \cdot \frac{\varepsilon^{1/2}\,d\varepsilon}{e^{\alpha + (\varepsilon/kT)} + 1} \qquad \ldots(6)$$

We introduce the *Fermi-distribution function* $f(\varepsilon)$ defined by

$$f(\varepsilon) = \frac{n(\varepsilon)}{g(\varepsilon)} = \frac{1}{e^{\alpha + (\varepsilon/kT)} + 1} = \frac{1}{e^{(\varepsilon - \varepsilon_F)/kT} + 1} \qquad \ldots(7)$$

where α in terms of *Fermi-energy* $\varepsilon_F(T)$ *at temperature* T is given by

$$\alpha = -\frac{\varepsilon_F(T)}{kT} \qquad \text{...(8)}$$

Let us consider the situation at absolute zero of temperature. At $T = 0$, $(\varepsilon - \varepsilon_F)/kT \to -\infty$ (when $\varepsilon < \varepsilon_F$) or $+\infty$ (when $\varepsilon > \varepsilon_F$). Thus at $T = 0$

$$f(\varepsilon) = \frac{1}{e^{-\infty}+1} = 1 \text{ for } \varepsilon < \varepsilon_F(0)$$

$$= \frac{1}{e^{\infty}+1} = 0 \text{ for } \varepsilon > \varepsilon_F(0) \qquad \text{...(9)}$$

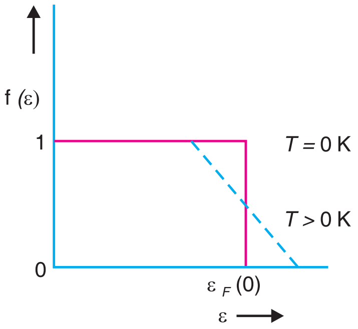


Fig. 52.3

Fig. 52.3 shows plot of Fermi-Dirac distribution function at $T = 0$K (solid line) and at low temperature $(kT/\varepsilon_F) << 1$ (dotted line).

At any temperature T,

$$f(\varepsilon) = 1/2 \text{ for } \varepsilon = \varepsilon_F.$$

The value of ε_F is determined by the condition that the total number of particles is constant at a given temperature T.

The expression for the number of fermions with energy between ε and $\varepsilon + d\varepsilon$ is written as

$$n(\varepsilon)\,d\varepsilon = f(\varepsilon)\,g(\varepsilon)\,d\varepsilon = \frac{8\pi m V}{h^3}\sqrt{2m}\,.\frac{\varepsilon^{1/2}\,d\varepsilon}{e^{(\varepsilon-\varepsilon_F)/kT}+1} \qquad \text{...(10)}$$

This expression is called *Fermi-Dirac energy distribution law.*

Thus at absolute zero all states with $0 < \varepsilon \leq \varepsilon_F$ (0) are completely filled and states with $\varepsilon > \varepsilon_F(0)$ are empty. The value $\varepsilon_F(0)$ of ε_F at $T = 0$K is determined by

$$N = \int_0^{\varepsilon_F(0)} n(\varepsilon)\,d\varepsilon = \int_0^{\varepsilon_F(0)} f(\varepsilon)\,g(\varepsilon)\,d\varepsilon = \int_0^{\varepsilon_F(0)} g(\varepsilon)\,d\varepsilon$$

$$= \int_0^{\varepsilon_F(0)} \frac{8\pi m V}{h^3}(2m)^{1/2}\varepsilon^{1/2}\,d\varepsilon = \frac{8\pi m V}{h^3}(2m)^{1/2}.\frac{2}{3}[\varepsilon_F(0)]^{3/2}$$

i.e.,

$$\varepsilon_F(0) = \frac{h^2}{2m}\left[\frac{3N}{8\pi V}\right]^{2/3} = \frac{h^2}{8m}\left[\frac{3N}{\pi V}\right]^{2/3} \qquad \text{...(11)}$$

$$\therefore \qquad \varepsilon_F(0) = (3\pi^2)^{2/3}\frac{\hbar^2}{2m}\left(\frac{N}{V}\right)^{2/3}$$

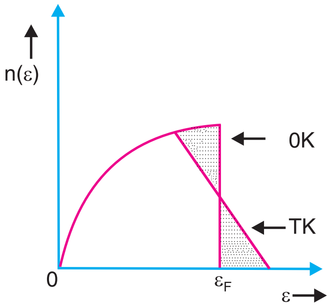


Fig. 52.4

The plot of $n(\varepsilon)$ against ε at $T = 0$ is shown in Fig. 52.4. After rising along a parabola from the origin, it drops abruptly to zero at $\varepsilon = \varepsilon_F$.

A typical plot of $n(\varepsilon)$ against ε at a much higher temperature $T = 1200$ K is also shown in Fig. 52.4. Since at any temperature the area under the corresponding curve gives the total number of electrons in the metal, the two shaded areas must be equal. It is seen that *even a considerable rise in temperature changes the energy state of only a very small fraction of the electrons.*

Let us now define a temperature T_F called *Fermi-temperature*, by

$$T_F = \frac{\varepsilon_F(0)}{k} = (3\pi^2)^{2/3} \frac{\hbar^2}{2mk}\left(\frac{N}{V}\right)^{2/3}$$

EXAMPLE 1. *Show that the mean energy of an electron in an electron gas at absolute zero temperature is* $\frac{3}{5}\varepsilon_F(0)$ *where* $\varepsilon_F(0)$ *is Fermi energy at absolute zero.* (*Rohilkhand,* 1993)

SOLUTION. The number of electrons in an electron gas having energies between ε and ε + *d*ε is given by

$$n(\varepsilon)\,d\varepsilon = \frac{8\pi mV}{h^3}\sqrt{(2m)}.\frac{\varepsilon^{1/2}\,d\varepsilon}{e^{(\varepsilon-\varepsilon_F)/kT}+1} \qquad ...(1)$$

where *m* is the mass of electron and *V* is the volume of the electron gas.

If *N* is the total number of electrons, then the Fermi energy ε_F is given by

$$\varepsilon_F(0) = \frac{h^2}{8m}\left(\frac{3N}{\pi V}\right)^{2/3}$$

$$\therefore \qquad \frac{V}{h^3} = \frac{3N}{8\pi m\sqrt{(2m)}}[\varepsilon_F(0)]^{-3/2}$$

Substituting this in Eq. (1), we get

$$n(\varepsilon)\,d\varepsilon = \frac{3}{2}N[\varepsilon_F(0)]^{-3/2}\frac{\varepsilon^{1/2}\,d\varepsilon}{e^{(\varepsilon-\varepsilon_F)/kT}+1}$$

At $T = 0$, all of the electrons have energies less than or equal to $\varepsilon_F(0)$, (*i.e.*, $\varepsilon \le \varepsilon_F$) so that at $T = 0$, we have

$$e^{(\varepsilon-\varepsilon_F)/kT} = e^{-\infty} = 0.$$

∴ At absolute zero

$$n(\varepsilon)\,d\varepsilon = \frac{3N}{2}[\varepsilon_F(0)]^{-3/2}\varepsilon^{1/2}\,d\varepsilon \qquad ...(2)$$

The total energy at absolute zero

$$E_0 = \int_0^{\varepsilon_F}\varepsilon\,n(\varepsilon)\,d\varepsilon$$

$$= \frac{3N}{2}[\varepsilon_F(0)]^{-3/2}\int_0^{\varepsilon_F(0)}\varepsilon^{3/2}\,d\varepsilon = \frac{3}{5}N\varepsilon_F(0)$$

∴ *Average energy* $\bar{\varepsilon}_0$ *at absolute zero is*

$$\bar{\varepsilon}_0 = \frac{E_0}{N} = \frac{3}{5}\varepsilon_F(0).$$

EXAMPLE 2. *Find the Fermi energy in copper on the assumption that each copper atom contributes one free electron to the electron gas. The density of copper is* 8.94×10^3 *kg / m*^{3} *and its atomic mass is* 63.5 *u.*

SOLUTION. Mass of one copper atom

$$= (63.5\text{ u})(1.66\times10^{-27}\text{ kg / u})$$

$$= 1.054\times10^{-25}\text{ kg}.$$

$$\frac{N}{V} = \frac{8.94\times10^3\ kg/m^3}{1.054\times10^{-25}\text{ kg}} = 8.48\times10^{28}\text{ atoms/m}^3.$$

$$= 8.48\times10^{28}\text{ electrons / m}^3.$$

The Fermi energy in copper is

$$\varepsilon_F = (3\pi^2)^{2/3}\frac{\hbar^2}{2m}\left(\frac{N}{V}\right)^{2/3} = (3\pi^2)^{2/3}\frac{(1.05\times10^{-34}\ \text{Js})^2}{2(9.11\times10^{-31}\ \text{kg})}(8.48\times10^{28}/\text{m}^3)^{2/3}$$

$$= 1.13\times10^{-18}\ \text{J} = 7.04\ \text{eV}.$$

52.3 Sommerfeld Theory of Electrical Conductivity

We consider that an electric field E is applied to the metal. The force upon an electron is $-eE$. The Boltzmann transport equation in the relaxation time approximation may be written as

$$e\boldsymbol{E}.\nabla_p f = \frac{f-f_0}{\tau} \qquad ...(1)$$

If we suppose that $\boldsymbol{E}$ is in the direction of x-axis, then the current will also be in the x-direction and we would be concerned with p_x — the x-component of the momentum $\boldsymbol{p}$ only. The above equation may be written as

$$e\boldsymbol{E}.\tau\frac{\partial f_0}{\partial p_x} = f-f_0 \qquad ...(2)$$

In writing this equation the difference between f and f_0 has been assumed to be quite small, so that $\frac{\partial f}{\partial p_x}$ has been approximated by $\frac{\partial f_0}{\partial p_x}$.

The kinetic energy of the electron is given by

$$\varepsilon = \left(p_x^2+p_y^2+p_z^2\right)\big/2m$$

So we can rewrite Eq. (2) as

$$v_x\ eE_x\tau\frac{\partial f_0}{\partial\varepsilon} = f-f_0. \qquad ...(3)$$

Now, we calculate the electrical current density J_x. We know that the number of electronic states per unit volume associated with an element $dp_x\ dp_y\ dp_z$ in momentum space is (including spin)

$$\frac{2}{h^3}dp_x\,dp_y\,dp_z.$$

Further, as f_0 is the electron distribution function (in terms of the total momentum p) when the system is in thermal equilibrium and in the absence of fields, the number

$$\frac{2}{h^3}f_0(\boldsymbol{p})\,dp_x\,dp_y\,dp_z$$

of these states is occupied. Now note that, in view of f_0 being the Fermi-Dirac distribution function, this number also represents the number of electrons which in equilibrium have their momenta in the range $dp_x\ dp_y\ dp_z$. Similarly, it can be argued that

The Electrical Conductivity Meter is being Demonstrated along with other Supplies for Calibrating and Sampling Liquid Manure.

$$\frac{2}{h^3} f(\boldsymbol{p})\, dp_x\, dp_y\, dp_z$$

represents the number of electrons which have their momenta in the same range $dp_x\, dp_y\, dp_z$ when the electric field is applied. f is here the field-value of the Fermi-Dirac distribution function. The current density is, therefore,

$$J_x = -\frac{2e}{h^3} \iiint v_x (f - f_0)\, dp_x\, dp_y\, dp_z \qquad \text{...(4)}$$

Note that the expression for J_x is simply the generalization of $J = -nev_D$.

Substituting the value of $(f - f_0)$ from Eq. (3) in Eq. (4), we get

$$J_x = -\frac{2e^2}{h^3} E_x \iiint v_x^2\, \tau \left(\frac{\partial f_0}{\partial \in} \right) dp_x\, dp_y\, dp_z \qquad \text{...(5)}$$

Now, τ is a function only of the energy and not of the direction of motion. $\dfrac{\partial f_0}{\partial \in}$ is also a function of ε alone. So the triple integral in Eq. (5) may be transformed into a single integral by replacing v_x^2 by $v^2/3$ and $dp_x\, dp_y\, dp_z$ by $4\pi p^2\, dp$. Expressing the integrand in terms ε alone, Eq. (5) becomes,

$$J_x = -\frac{16\pi e^2 (2m)^{1/2}}{3h^3} E_x \int_0^\infty \varepsilon^{3/2}\, \tau(\varepsilon) \left(\frac{\partial f_0}{\partial \varepsilon} \right) d\varepsilon, \qquad \text{...(6)}$$

It must now be noted that $\dfrac{\partial f_0}{\partial \varepsilon}$ has an appreciable value only in an energy range of a few kT about the Fermi energy ε_F. There will thus be no significant value of the integral in the above equation except when ε takes on values within a few kT units of ε_F. If $\varepsilon^{3/2}\, \tau(\varepsilon)$ does not vary much over this energy range, then to a good approximation, it may be replaced by the quantity $\varepsilon_F^{3/2}\, \tau(\varepsilon_f)$ and may be taken outside the integral. Furthermore

$$\int_0^\infty \left(\frac{\partial f_0}{\partial \varepsilon} \right) d\varepsilon = -1.$$

So, Eq. (6) becomes,

$$J_x = \frac{16\pi e^2 (2m)^{1/2}}{3h^3} E_x\, \varepsilon_F^{3/2}\, \tau(\varepsilon_f)$$

The Fermi energy of a metal is given by

$$\varepsilon_F = \frac{h^2}{2m} \left(\frac{3N}{8\pi V} \right)^{2/3}$$

where m is the mass of electron and N is the number of free electrons in a volume V of the metal.

$$\therefore \qquad J_x = \frac{ne^2 \tau(\varepsilon_f)}{m} E_x.$$

Here n is the number of electrons per unit volume. $\tau(\varepsilon_f)$ is collision time.

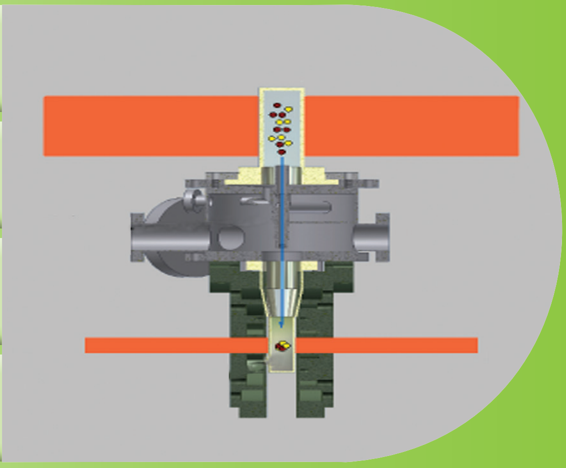
Electrical conductivity

$$\sigma = \frac{J_x}{E_x} = \frac{ne^2 \tau(\varepsilon_f)}{m} \qquad \text{...(7)}$$

The result (Eq. 7) obtained is essentially the same as obtained by Drude-Lorentz using the Maxwell-Boltzmann distribution except that $\tau(v)$ has been replaced by $\tau(\varepsilon_f)$.

CHAPTER

53



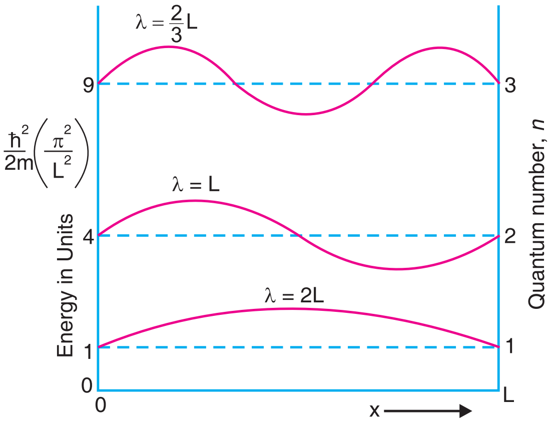
FREE ELECTRON FERMI GAS

AT A GLANCE

53.1	Energy Levels in one Dimension	**53.2**	Free Electron Gas in Three Dimensions

53.1 Energy Levels in One Dimension

By a free electron Fermi gas, we mean a gas of free electrons subject to the Pauli principle. Consider a free electron gas in one dimension. An electron of mass m is confined to a length L by infinite barriers (Fig. 53.1).



λ = 2/3 L
9
3
ħ²/2m (π²/L²)
λ = L
4
2
λ = 2L
Energy in Units
1
1
0
0
L
x
Quantum number, n

Fig. 53.1

The wave function ψn (x) of the electron is a solution of the Schrodinger equation

$$H\psi = \varepsilon\psi \qquad \text{...(1)}$$

Hamiltonian H = Kinetic energy + potential energy.

The potential energy is neglected. So we have

$$H = p^2/2m \qquad \text{...(2)}$$

Here, p is the momentum. In quantum theory p is represented by $-i\hbar\frac{d}{dx}$.

$$\therefore \qquad H\psi_n = -\frac{\hbar^2}{2m}\frac{d^2\psi_n}{dx^2} = \varepsilon_n\psi_n, \qquad \text{...(3)}$$

Here, ε_n is the energy of the electron in the orbital.

The boundary conditions are $\psi_n(0) = 0$; $\psi_n(L) = 0$, as imposed by the infinite potential energy barriers.

They are satisfied if the wavefunction is sinelike with an integral number n of half-wavelengths between 0 and L.

$$\therefore \qquad \psi_n = A\sin\left(\frac{2\pi}{\lambda_n}x\right); \quad \frac{1}{2}n\lambda_n = L, \qquad \text{...(4)}$$

or

$$\psi_n = A\sin\left(\frac{n\pi x}{L}\right) \quad \text{where } A \text{ is a constant.}$$

Thus, only for integral values of n there are allowed wave functions $\psi_n(x)$ and the corresponding energy values ε_n.

The energy ε_n is given by

$$\varepsilon_n = \frac{\hbar^2}{2m}\left(\frac{n\pi}{L}\right)^2 \qquad \text{...(5)}$$

Fig. 53.1 shows first three energy levels and wave functions of a free electron of mass m confined to a line of length L.

We now discuss the distribution of electrons among the allowed energy levels and determine the Fermi energy, total energy and the density of states.

(i) Fermi Energy

We want to accommodate N electrons on the line. According to the *Pauli exclusion principle*, no two electrons can have all their quantum numbers identical. In a linear solid the quantum numbers of a conduction electron orbital are n and m_s. n is any positive integer and $m_s = \pm\frac{1}{2}$, according to the spin orientation. Each energy level with quantum number n can accommodate two electrons, one with spin up and one with spin down.

Suppose there are six electrons. Then in the ground state of the system the filled orbitals are those given in the table.

n	m_s	*electron occupancy*	n	m_s	*electron occupancy*
1	↑	1	3	↑	1
1	↓	1	3	↓	1
2	↑	1	4	↑	0
2	↓	1	4	↓	0

More than one orbital may have the same energy. The number of orbitals with the same energy is called the *degeneracy.*

Let n_F denote the topmost filled energy level. Let N be the total number of electrons to be accommodated. Then for even N, $2n_F = N$ determines n_F.

The *Fermi energy* ε_F is defined as the energy of the topmost filled level in the ground state of the N electron system. For $n = n_F$, Eq. (5) gives

$$\varepsilon_F = \frac{\hbar^2}{2m}\left(\frac{n_F\,\pi}{L}\right)^2 = \frac{\hbar^2}{2m}\left(\frac{N\pi}{2L}\right)^2 \qquad \text{...(6)}$$

(ii) Total Energy

The total energy, ε_0, of all the N electrons in the ground state is determined by summing up the energies of the individual electrons. For N electrons, the number of filled energy levels is $N/2$.

$$\therefore \qquad \varepsilon_0 = 2\sum_{n=1}^{N/2}\varepsilon_n = 2\frac{\hbar^2}{2m}\left(\frac{\pi}{L}\right)^2\sum_{n=1}^{N/2}n^2$$

Here the factor of 2 appears because each level contains two electrons with equal energies.

$$\text{Now,} \qquad \sum_{n=1}^{N/2}n^2 = 1^2 + 2^2 + 3^2 + \ldots\ldots + S^2 \quad \text{where } S = N/2.$$

$$= \frac{1}{6}S\,(2S^2 + 3S + 1) \cong \frac{1}{3}S^3 \text{ for } S \gg 1$$

$$\cong \frac{1}{3}(N/2)^3$$

$$\varepsilon_0 = 2\frac{\hbar^2}{2m}\left(\frac{\pi}{L}\right)^2\frac{1}{3}\left(\frac{N}{2}\right)^3 = \frac{1}{3}\frac{\hbar^2}{2m}\left(\frac{N\pi}{2L}\right)^2 N = \frac{1}{3}N\varepsilon_F \qquad \text{...(7)}$$

(iii) Density of States

The density of states, $D(\varepsilon)$, is defined as the number of electronic states present in a unit energy range.

$$D(\varepsilon) = \frac{dn}{d\varepsilon} \qquad \text{...(8)}$$

For a free electron gas, since each energy level contains two electronic states, one with spin up and the other with spin down, the actual density of states is twice the value given by Eq. (8).

$$D(\varepsilon) = 2\frac{dn}{d\varepsilon}$$

Differentiating Eq. (5), we get

$$d\varepsilon = \frac{\hbar^2}{2m}\left(\frac{\pi}{L}\right)^2 2\,n\,dn.$$

$$D(\varepsilon) = \frac{2dn}{d\varepsilon} = 2\cdot\frac{2m}{\hbar^2}\left(\frac{L}{\pi}\right)^2\frac{1}{2n} = \frac{4L}{h}\left(\frac{m}{2\varepsilon_n}\right)^{1/2}$$

The plot of $D(\varepsilon)$ versus ε is shown in Fig. 53.2.

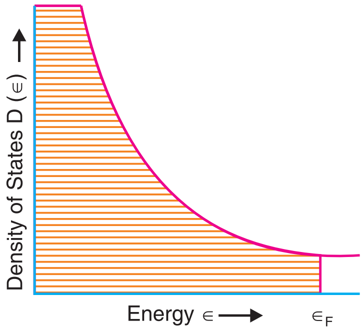


Fig. 53.2

53.2 Free Electron Gas in Three Dimensions

The free-particle Schrodinger equation in three dimensions is

$$-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right)\psi_k(\mathbf{r}) = \varepsilon_k\,\psi_k(\mathbf{r}) \qquad \text{...(1)}$$

Suppose the electrons are confined to a cube of edge L.

The wavefunction is the standing wave

$$\psi_n(\mathbf{r}) = A \sin(\pi n_x x / L) \sin(\pi n_y y / L) \sin(\pi n_z z / L), \qquad \text{...(2)}$$

Here, n_x, n_y, n_z are positive integers. The origin is at one corner of the cube.

We now require the wavefunctions to be periodic in x, y, z with period L.

Thus

$$\begin{aligned} \psi(x+L, y, z) &= \psi(x, y, z), \\ \psi(x, y+L, z) &= \psi(x, y, z), \\ \psi(x, y, z+L) &= \psi(x, y, z). \end{aligned} \qquad \text{...(3)}$$

Wavefunctions satisfying the free-particle Schrodinger equation and the periodicity condition are of the form of a travelling plane wave.

$$\psi_k(\mathbf{r}) = \exp(i\,\mathbf{k}.\,\mathbf{r}) = \exp i(k_x x + k_y y + k_z z) \qquad \text{...(4)}$$

The components of the wavevector $\mathbf{k}$ (k_x, k_y, k_z) must satisfy

$$k_x = 0, \pm\frac{2\pi}{L}, \pm\frac{4\pi}{L}, \ldots\ldots\ldots \qquad \text{...(5)}$$

Similar results are obtained for k_y and k_z.

Any component of $\mathbf{k}$ is of the form $2n\pi / L$, where n is a positive or negative integer.

The quantum number m_s represents the spin direction. Thus the state of an electron is specified completely by a set of four quantum numbers k_x, k_y, k_z and m_s.

On substituting Eq. (4) in Eq. (1), we have the energy ε_k of the orbital with wave vector $\mathbf{k}$.

$$\varepsilon_k = \frac{\hbar^2}{2m}k^2 = \frac{\hbar^2}{2m}(k_x^2 + k_y^2 + k_z^2). \qquad \text{...(6)}$$

The magnitude of the wavevector $\mathbf{k}$ is related to the wavelength λ by

$$k = \frac{2\pi}{\lambda} \qquad \text{...(7)}$$

Thus the energy spectrum consists of discrete energy levels.

(a) The Fermi Energy

In quantum mechanics, the linear momentum $\mathbf{p}$ is represented by the operator $\mathbf{p} = -i\hbar\nabla$. When operated on Eq. (4),

$$\mathbf{p}\,\psi_k(\mathbf{r}) = -i\hbar\nabla\,\psi_k(\mathbf{r}) = \hbar\,\mathbf{k}\,\psi_k(\mathbf{r}). \qquad \text{...(8)}$$

This indicates that the plane wave ψ_k is an eigenfunction of the linear momentum with the eigenvalue $\hbar\mathbf{k}$. The particle velocity in the orbital $\mathbf{k}$ is given by $\mathbf{v} = \hbar\mathbf{k} / m$.

In the ground state of a system of N free electrons, the occupied orbitals may be represented as points inside a sphere in $\mathbf{k}$ space. The energy at the surface of the sphere is the *Fermi energy.*

Let $\mathbf{k_F}$ be the wave vector from the origin of the $\mathbf{k}$-space to the surface of the sphere (Fig. 53.3). Then, using Eq. (6), the Fermi energy is written as

$$\varepsilon_F = \frac{\hbar^2}{2m}k_F^2 \qquad \text{...(9)}$$

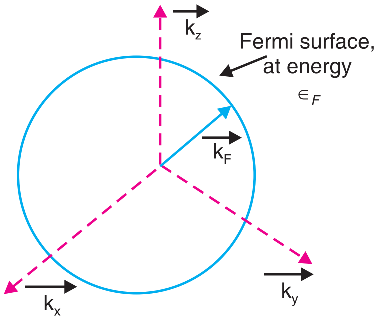


Fig. 53.3

From Eq. (5) we see that there is one allowed wavevector or one distinct triplet of quantum numbers k_x, k_y, k_z which corresponds to the volume element $(2\pi / L)^3$ of **k** space. Thus, in the sphere called the *Fermi sphere* of volume $(4\pi / 3)\, k^3{}_F$, the total number of orbitals is

$$2 \cdot \frac{4\pi k_F^3 / 3}{(2\pi / L)^3} = \frac{V}{3\pi^2} k_F^3 = N. \qquad ...(10)$$

Here, the factor 2 appears because there are two allowed values of m_s for each allowed value of **k**.

From Eq. (10),
$$k_F = \left(\frac{3\pi^2 N}{V}\right)^{1/3}. \qquad ...(11)$$

This shows that the value of k_F depends on the concentration of electrons, N / V, and is independent of the mass of the electron.

Using Eq. (9),
$$\varepsilon_F = \frac{\hbar^2}{2m}\left(\frac{3\pi^2 N}{V}\right)^{2/3} \qquad ...(12)$$

The Fermi energy depends on both the electronic concentration and mass.

The electron velocity, v_F, at the Fermi surface is

$$v_F = \frac{\hbar k_F}{m} = \frac{\hbar}{m}\left(\frac{3\pi^2 N}{V}\right)^{1/3} \qquad ...(13)$$

(b) Density of States

The number of orbitals per unit energy range, $D(\varepsilon)$, is called the *density of states*. We use Eq. (12) for the total number of orbitals of energy $\leq \varepsilon$.

$$N = \frac{V}{3\pi^2}\left(\frac{2m\varepsilon}{\hbar^2}\right)^{3/2} \qquad ...(14)$$

The density of states is

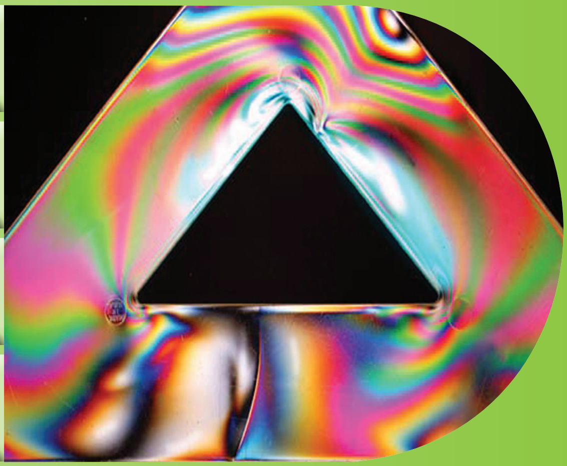
$$D(\varepsilon) \equiv \frac{dN}{d\varepsilon} = \frac{V}{2\pi^2} \cdot \left(\frac{2m}{\hbar^2}\right)^{3/2} \cdot \varepsilon^{1/2} \qquad ...(15)$$

EXERCISE

1. Obtain expressions for the Fermi energy, the total energy and the density of states for a free electron gas in one dimension.
2. Derive an expression for the Fermi energy of a free electron gas in three dimensions. (*Raj*., 1975)

CHAPTER

54



DIELECTRIC PROPERTIES OF SOLIDS

AT A GLANCE

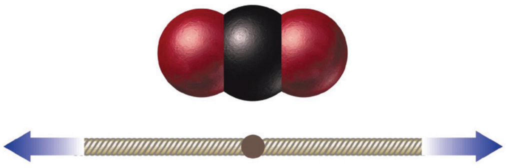
54.1	Introduction	**54.2**	Polarization
54.3	Macroscopic Electric Field	**54.4**	Depolarization Field
54.5	Local Electric Field at an Atom	**54.6**	Dielectric Constant
54.7	Polarizability	**54.8**	Derivation of the Clausius-Mossotti Relation
54.9	Frequency Dependence of Total Polarizability	**54.10**	Electronic Polarizability
54.11	Classical Theory of Electronic Polarizability		

54.1 Introduction

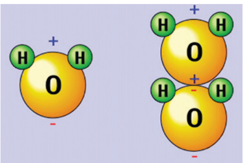
Dielectric. *A dielectric* (*or an insulator*) *is a material in which all the electrons are tightly bound to the nuclei of the atoms*. Thus there are no free electrons to carry current. Hence the electrical conductivity of a dielectric is very low. (The conductivity of an ideal dielectric is zero). Glass, plastic, mica, oil, etc., are examples of dielectrics.

Non-polar and Polar Molecules— The molecules of dielectrics may be classified as 'nonpolar' and 'polar'.

A **non-polar molecule** is one in which the centre of gravity of the positive charges (protons) coincides with the centre of gravity of the negative charges (electrons). The non-polar



Non-polar Molecules (CO_2).



Polar Molecules (H_2O).

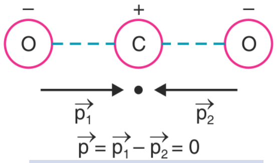


Fig. 54.1

molecules have symmetrical structure and zero electric dipole moment. Examples are: H_2, N_2, O_2, CO_2, benzene, etc (Fig. 54.1).

A polar molecule is one in which the centre of gravity of the positive charges is separated from the centre of gravity of the negative charges by a finite distance. The polar molecule is thus an electric dipole and has an intrinsic permanent dipole moment. Examples are: H_2O, HCl, CO, N_2O, NH_3, etc., (Fig. 54.2). The permanent dipole moment of a molecule of water has the magnitude $p = 6.3 \times 10^{-30}$ Cm and is directed from O^{--} ion toward the midpoint of the line connecting the H^+ ions. The total dipole moment is defined as

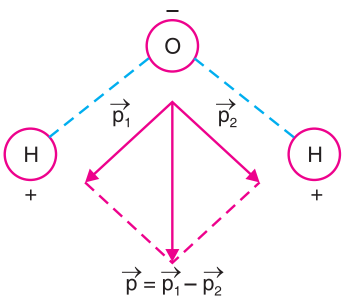


Fig. 54.2

$$\mathbf{p} = \Sigma q_n \mathbf{r}_n, \quad \text{...(1)}$$

Here, $\mathbf{r}_n$ is the position vector of the charge q_n. The value of the sum will be independent of the origin chosen for the position vectors, provided that the system is neutral.

The electric field at a point **r** from a dipole moment **p** is given by

$$\mathbf{E}(\mathbf{r}) = \frac{3(\mathbf{p}.\mathbf{r})\mathbf{r} - r^2\mathbf{p}}{4\pi\epsilon_0 r^5} \quad \text{...(2)}$$

54.2 Polarization

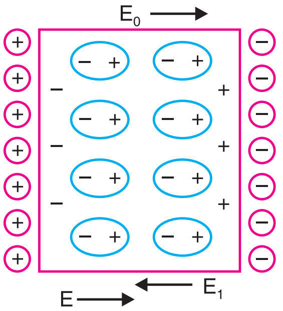


Fig. 54.3

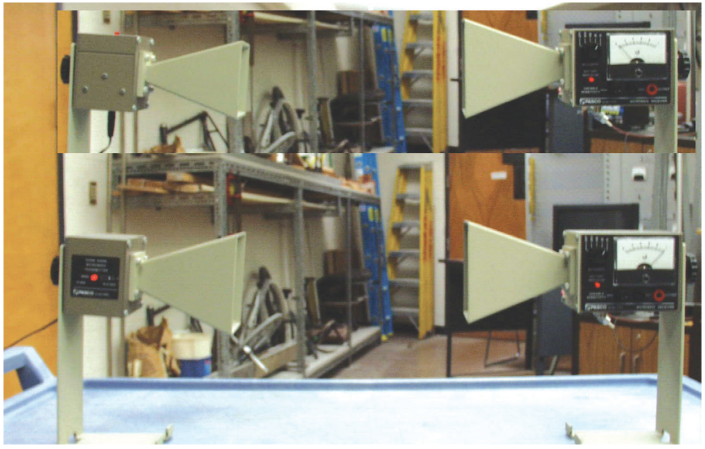
When a dielectric is placed in an external electric field $\mathbf{E}_0$ (Fig. 54.3), the positive and negative charges are displaced from their equilibrium positions by very small distances throughout the volume of the dielectric. This results in the formation of a large number of dipoles each having some dipole moment in the direction of the field. The material is said to be polarized with a *Polarization* **P**.

The polarization **P** is defined as the dipole moment per unit volume, averaged over the volume of a cell.

The effect of polarization is to reduce the magnitude of the external field $\mathbf{E}_0$. Thus the magnitude of the resultant field is less than the applied field, *i.e.*, $E < E_0$.

We may write

$$\mathbf{E} = \mathbf{E}_0 + \mathbf{E}_1 \quad \text{...(1)}$$



Polarization of Microwaves.

The field $\mathbf{E}_1$ is called the *depolarization field*, for within the body it tends to oppose the applied field $\mathbf{E}_0$. For ordinary electric fields, the polarization **P** is proportional to the macroscopic field **E**. It is expressed as

$$\mathbf{P} = \varepsilon_0 \chi_e \mathbf{E} \qquad ...(2)$$

Here, χ_e is the electric susceptibility.

The electric displacement vector **D** for an isotropic or cubic medium is defined as

$$\mathbf{D} = \varepsilon_0 \varepsilon_r \mathbf{E} = \varepsilon_0 \mathbf{E} + \mathbf{P} \qquad ...(3)$$

Here, ε_r is the relative permittivity or dielectric constant of the dielectric.

ε_r is related to the electric susceptibility as

$$\varepsilon_r = 1 + \chi_e \qquad ...(4)$$

The physical quantities of primary interest are the field vectors **E** and **D**, the polarization **P**, together with the electric susceptibility χ_e and dielectric constant ε_r.

54.3 Macroscopic Electric Field

One contribution to the electric field inside a body is that of the applied electric field. It is defined as

$\mathbf{E}_0 \equiv$ field produced by fixed charges external to the body.

The other contribution to the electric field is the sum of the fields of all charges that constitute the body. If the body is neutral, the contribution to the average field may be expressed in terms of the sum of the fields of atomic dipoles.

The macroscopic field is a volume average over a region large compared with atomic dimensions. We define the average electric field $\mathbf{E}(\mathbf{r}_0)$ as the average field over the volume of the crystal cell that contains the lattice point $\mathbf{r}_0$.

$$\mathbf{E}(\mathbf{r}_0) = \frac{1}{V_c}\int dV\, \mathbf{e}(\mathbf{r}). \qquad ...(1)$$

Here, **e (r) is the microscopic electric field at the point r.**

We call **E** the macroscopic electric field. To find the contribution of the polarization to the macroscopic field, we can simplify the sum over all the dipoles in the specimen. By a theorem of electrostatics the macroscopic electric field caused by a uniform polarization **P** is equal to the electric field in vacuum of a fictitious surface charge density $\sigma = \hat{n}.\mathbf{P}$ on the surface of the body. Here $\hat{n}$ is the unit normal to the surface, drawn outward from the polarized matter.

Let us apply the result to a thin dielectric slab (Fig. 54.4) with a uniform volume polarization **P**. The electric field $\mathbf{E}_1(\mathbf{r})$ produced by the polarization is equal to the field produced by the fictitious surface charge density $\sigma = \hat{\mathbf{n}}.\mathbf{P}$ on the surface of the slab. On the upper boundary the unit vector $\hat{\mathbf{n}}$ is directed upward and on the lower boundary $\hat{\mathbf{n}}$ is directed downward. The upper boundary bears the fictitious charge $\sigma = \hat{\mathbf{n}}.\mathbf{P} = P$ per unit area. The lower boundary bears $-P$ per unit area.

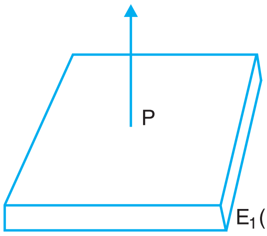


Fig. 54.4

By Gauss's law, the electric field E_1 due to these charges at any point between the plates is given by

$$E_1 = \frac{P}{\epsilon_0} \qquad ...(2)$$

The total macroscopic field inside the slab is

$$\mathbf{E} = \mathbf{E}_0 + \mathbf{E}_1 = \mathbf{E}_0 - \frac{P}{\epsilon_0}\hat{\mathbf{z}}$$

Here, $\hat{\mathbf{z}}$ is the unit vector normal to the plane of the slab.

We define

$\mathbf{E}_1 \equiv$ field of the surface charge density $\hat{\mathbf{n}} \cdot \mathbf{P}$ on the boundary.

54.4 Depolarization Field

Let us determine the macroscopic field **E** within a dielectric that has prolate ellipsoidal form. In ellipsoids, a uniform polarization produces a uniform depolarization field inside the body. Then, the only contributions to the macroscopic field are from $\mathbf{E}_0$ and $\mathbf{E}_1$:

$$\mathbf{E} = \mathbf{E}_0 + \mathbf{E}_1 \qquad ...(1)$$

Here $\mathbf{E}_0$ is the applied field and $\mathbf{E}_1$ is the field due to the uniform polarization.

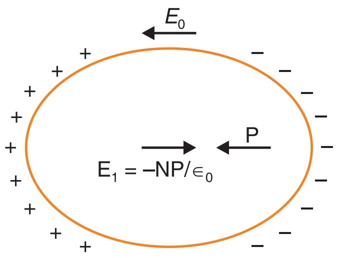


Fig. 54.5

The polarization **P** produces a surface distribution of charge whose polarity gives rise to a field within the dielectric that is directed oppositely to the applied field $\mathbf{E}_0$ (Fig. 54.5). This is the depolarizing field $\mathbf{E}_1$.

Let P_x, P_y, P_z be the components of the polarisation **P** along the principal axes of the ellipsoid. Then the components of the depolarization field $\mathbf{E}_1$ are given by

$$E_{1x} = -\frac{N_x P_x}{\varepsilon_0}\,;\; E_{1y} = -\frac{N_y P_y}{\varepsilon_0}\,;\, E_{1z} = -\frac{N_z P_z}{\varepsilon_0} \qquad ...(2)$$

Here N_x, N_y, N_z are the depolarization factors.

Their values depend on the ratios of the principal axes of the ellipsoid.

Also N_x, N_y, N_z have positive values and $N_x + N_y + N_z = 1$.

If $\mathbf{E}_0$ is uniform and parallel to a principal axis of the ellipsoid, then $\mathbf{E}_1 = -\frac{N\mathbf{P}}{\varepsilon_0}$.

We write the macroscopic field within the dielectric as

$$\mathbf{E} = \mathbf{E}_0 - \frac{N\mathbf{P}}{\varepsilon_0}, \qquad ...(3)$$

N is the geometrical depolarizing factor.

In limiting cases *N* has the values:

Shape	*Axis*	*N*
Sphere	any	1/3
Thin slab	in plane	0
Thin slab	normal	**1**
Long circular cylinder	longitudinal	0
Long circular cylinder	transverse	1/2

Depolarization factor *N* for three important geometrical shapes is shown in Fig. 54.6.

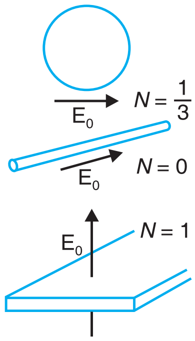


Fig. 54.6

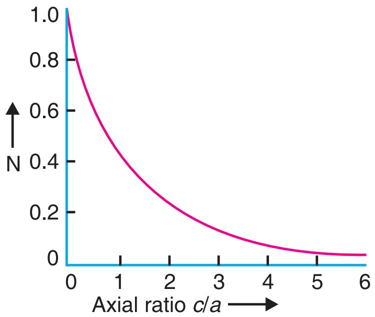


Fig. 54.7

Values of N parallel to the figure axis of ellipsoids of revolution, as a function of the axial ratio c/a are plotted in Fig. 54.7.

A uniform applied field $\mathbf{E}_0$ will induce uniform polarization in an ellipsoid. We introduce the dielectric susceptibility χ_e such that the relation

$$\mathbf{P} = \varepsilon_0 \chi_e \mathbf{E} \qquad \text{...(4)}$$

connects the macroscopic field **E** inside the ellipsoid with the polarization **P**.

Substituting for **E** from Eq. (3), we get

$$\mathbf{P} = \frac{\chi_e \varepsilon_0}{1 + N\chi_e} \mathbf{E}_0$$

The value of the polarization depends on the depolarization factor N.

54.5 Local Electric Field at an Atom

The value of the local electric field $\mathbf{E}_{local}$ that acts at the site of an atom is not the same as the macroscopic average electric field **E**. Consider the local field at a site with a cubic arrangement of neighbours in a crystal of spherical shape. The macroscopic electric field in a sphere is

$$\mathbf{E} = \mathbf{E}_0 + \mathbf{E}_1 = \mathbf{E}_0 - \frac{1}{3\epsilon_0}\mathbf{P} \qquad \text{...(1)}$$

But consider the field that acts on the atom at the centre of the sphere. Let all the dipoles be parallel to the z axis and have magnitude p. Then, the z component of the field at the centre due to all other dipoles is,

$$E_{dipole} = \frac{p}{4\pi\epsilon_0} \sum_i \frac{3z_i^2 - r_i^2}{r_i^5} = \frac{p}{4\pi\epsilon_0} \sum_i \frac{2z_i^2 - x_i^2 - y_i^2}{r_i^5} \qquad \text{...(2)}$$

Here, p is the magnitude of the atomic dipole moment and i's refer to different atoms.

The x, y, z directions are equivalent because of the symmetry of the lattice and of the sphere. Thus

$$\sum_i \frac{z_i^2}{r_i^5} = \sum_i \frac{x_i^2}{r_i^5} = \sum_i \frac{y_i^2}{r_i^5}$$

$$\therefore \qquad E_{dipole} = 0.$$

The correct local field is just equal to the applied field, $\mathbf{E}_{local} = \mathbf{E}_0$, for an atom site with a cubic environment in a spherical specimen. Thus the local field is not the same as the macroscopic average field **E**.

We now develop an expression for the local field at a general lattice site, not necessarily of cubic symmetry. The local field at an atom is the sum of the electric field $\mathbf{E}_0$ from external sources and of the field from the dipoles within the specimen.

We write

$$\mathbf{E}_{local} = \mathbf{E}_0 + \mathbf{E}_1 + \mathbf{E}_2 + \mathbf{E}_3 \qquad ...(3)$$

Here,

$\mathbf{E}_0$ = field produced by fixed charges external to the body

$\mathbf{E}_1$ = depolarization field, from a surface charge density $\hat{\mathbf{n}}.\mathbf{P}$ on the outer surface of the specimen,

$\mathbf{E}_2$ = Lorentz cavity field: field from polarization charges on inside of a spherical cavity cut out of the specimen with the reference atom as centre (Fig. 54.8),

$\mathbf{E}_3$ = field of atoms inside the spherical cavity.

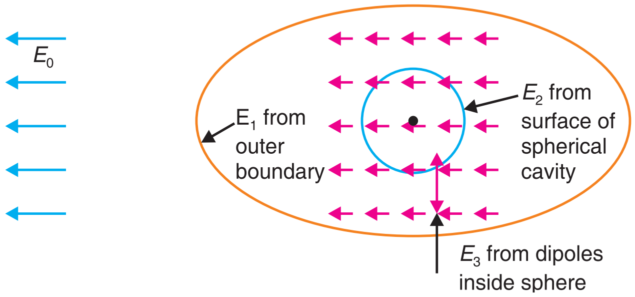

E_0
E_1 from outer boundary
E_2 from surface of spherical cavity
E_3 from dipoles inside sphere

Fig. 54.8

Let us now calculate the total local field acting on an atom in a cubic site of a dielectric.

(*i*) $\mathbf{E}_1$: The value of this field depends on the geometrical shape of the external surface.

For a sphere, $\mathbf{E}_1 = -\dfrac{1}{3\varepsilon_0}\mathbf{P}$.

(*ii*) **Lorentz Field**, $\mathbf{E}_2$

The field $\mathbf{E}_2$ due to the polarization charges on the surface of the fictitious cavity was calculated by Lorentz.

Let θ be the polar angle referred to the polarization direction (Fig. 54.9).

The surface charge density on the surface of the cavity is $-P\cos\theta$.

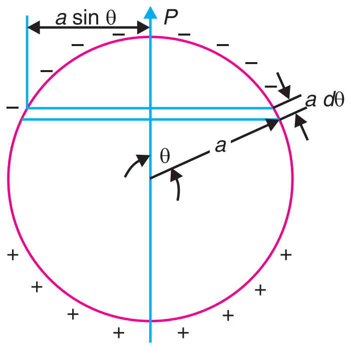


Fig. 54.9

The area of the ring-shaped element is $2\pi a^2 \sin\theta\, d\theta$.

Hence charge on this ring is $P\cos\theta \cdot 2\pi a^2 \sin\theta \cdot d\theta$.

The electric field at the centre of the spherical cavity of radius a is

$$\mathbf{E}_2 = \int_0^{\pi} \frac{1}{4\pi\epsilon_0}(a^{-2})(2\pi a^2 \sin\theta\, d\theta)(P\cos\theta)(\cos\theta) = \frac{P}{3\epsilon_0} \qquad ...(4)$$

This is the negative of the depolarization field $\mathbf{E}_1$ in a polarized sphere.

$\therefore \qquad \mathbf{E}_1 + \mathbf{E}_2 = 0$ for a sphere.

(*iii*) Field of Dipoles Inside Cavity, $\mathbf{E}_3$

The field $\mathbf{E}_3$ due to the dipoles within the spherical cavity is the only term that depends on the crystal structure. We showed for a reference site with cubic surroundings in a sphere that $\mathbf{E}_3 = 0$ if all the atoms may be replaced by point dipoles parallel to each other.

The total local field at a cubic site is, from Eqs. (3) and (4),

$$\mathbf{E}_{local} = \mathbf{E}_0 + \mathbf{E}_1 + \frac{\mathbf{P}}{3\in_0} = \mathbf{E} + \frac{1}{3\in_0}\mathbf{P}. \qquad ...(5)$$

Equation (5) is the **Lorentz relation.**

54.6 Dielectric Constant

The dielectric constant $\in_r$ of an isotropic or cubic medium relative to vacuum is defined in terms of the macroscopic field E.

The electric displacement vector for an isotropic or cubic medium can be defined as

$$\mathbf{D} = \varepsilon_0\varepsilon_r\mathbf{E} = \varepsilon_0\mathbf{E} + \mathbf{P} \qquad ...(1)$$

Here, ε_r is called the *relative permittivity* or *dielectric constant* of the dielectric.

$$\therefore \qquad \varepsilon_r = \frac{\varepsilon_0\,\mathbf{E} + \mathbf{P}}{\varepsilon_0\,\mathbf{E}} = 1 + \chi_e$$

The susceptibility is related to the dielectric constant by

$$\chi_e = \frac{\mathbf{P}}{\varepsilon_0\,\mathbf{E}} = \in_r - 1$$

Experimental Setup to Determine the Dielectric Constant.

54.7 Polarizability

The *polarizability*, α, of an atom is defined as the dipole moment per unit local electric field at the atom, *i.e.*,

$$p = \alpha\, E_{local}, \qquad ...(1)$$

Here, p is the dipole moment, $\mathbf{E}_{local}$ is the electric field acting at the site of the atom. We call this local field. The polarizability is an atomic property. But the dielectric constant is a macroscopic property which depends upon the arrangement of atoms within the crystal.

For a non-spherical atom α is a tensor.

If all the atoms have the same polarizability and there are N number of atoms per unit volume, the polarization can be expressed as

$$P = Np = N\alpha E_{local} \qquad ...(2)$$

The polarization of a crystal may be expressed approximately as the product of the polarizabilities of the atoms times the local electric field

$$P = \sum_j N_j\, p_j = \sum_j N_j\, \alpha_j\, E_{loc}\,(j) \qquad ...(3)$$

Here, N_j is the concentration and α_j the polarizability of atoms j, and $E_{loc}\,(j)$ is the local field at atom sites j.

54.8 Derivation of the Clausius-Mossotti Relation

The polarisation (P) of the dielectric is given by

$$P = \sum_j N_j p_j = \sum_j N_j \alpha_j E_{loc}(j) \qquad \text{...(1)}$$

Here, N_j is the concentration and α_j the polarizability of atoms j, and E_{loc} (j) is the local field at atom sites j. p_j is the dipole moment of the jth atom.

The total local field at a cubic site is given by the Lorentz relation

$$\mathbf{E}_{local} = \mathbf{E} + \frac{1}{3\epsilon_0}\mathbf{P} \qquad \text{...(2)}$$

Here, **E** is the macroscopic electric field.

$$P = (\Sigma N_j \alpha_j)\left(E + \frac{1}{3\varepsilon_0}P\right) \qquad \text{...(3)}$$

$$P\left(1 - \frac{\Sigma N_j \alpha_j}{3\varepsilon_0}\right) = E \Sigma N_j \alpha_j$$

Thus, the dielectric susceptibility is given by,

$$\chi_e = \frac{P}{\varepsilon_0 E} = \frac{\Sigma N_j \alpha_j}{\varepsilon_0\left(1 - \frac{\Sigma N_j \alpha_j}{3\varepsilon_0}\right)} \qquad \text{...(4)}$$

Further, the dielectric constant,

$$\varepsilon_r = 1 + \chi_e = 1 + \frac{\Sigma N_j \alpha_j}{\varepsilon_0\left(1 - \frac{\Sigma N_j \alpha_j}{3\varepsilon_0}\right)} = \frac{3\Sigma N_j \alpha_j}{3\varepsilon_0 - \Sigma N_j \alpha_j}$$

$$\frac{\varepsilon_r - 1}{\varepsilon_r + 2} = \frac{1}{3\varepsilon_0}\Sigma N_j \alpha_j \qquad \text{...(5)}$$

Equation (5) is the *Clausius-Mossotti relation.*

It relates the dielectric constant to the atomic polarizability provided the condition of cubic symmetry holds.

54.9 Frequency Dependence of Total Polarizability

The net polarizability of a dielectric material results from three main contributions—the electronic polarizability, ionic polarizability, and dipolar or orientational polarizability. The electronic contribution arises from the displacement of the electron shell relative to a nucleus. The ionic contribution comes from the displacement of a charged ion with respect to other ions. The dipolar polarizability arises from molecules with a permanent electric dipole moment that can change orientation in an applied electric field.

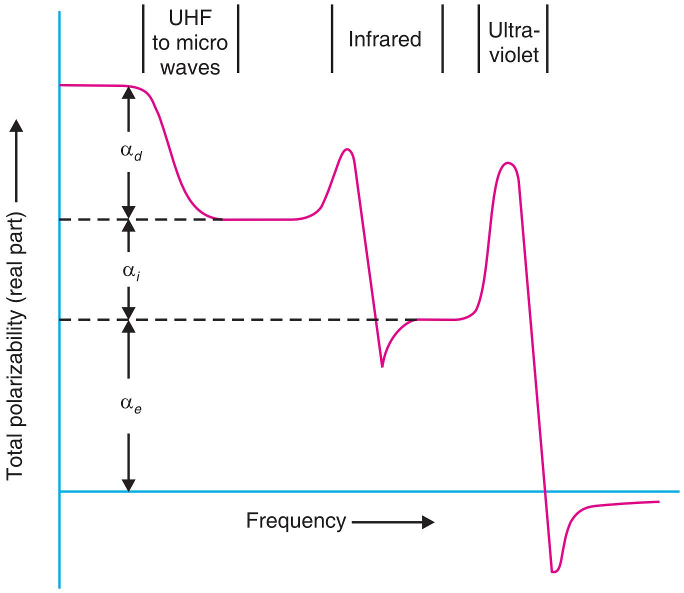


Fig. 54.10

The total polarizability of a dielectric is given by the expression:

$$\alpha = \alpha_e + \alpha_i + \alpha_d \quad \text{...(1)}$$

Figure 54.10 depicts the frequency dependence of the several contributions to the polarizability.

54.10 Electronic Polarizability

The electronic polarizability arises due to the displacement of the electron cloud of an atom relative to its nucleus in the presence of an applied electric field (Fig. 54.11). The dielectric constant at optical frequencies arises almost entirely from the electronic polarizability. The dipolar and ionic contributions are small at high frequencies because of the inertia of the molecules and ions.

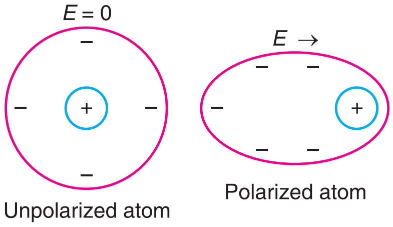


Fig. 54.11

At optical frequencies, $n^2 = \varepsilon_r$, where n is the refractive index.

The Clausius-Mossotti relation is written as

$$\frac{n^2-1}{n^2+2} = \frac{1}{3\varepsilon_0}\Sigma N_j \alpha_j \text{ (electronic).}$$

54.11 Classical Theory of Electronic Polarizability

An electron bound harmonically to an atom will show resonance at a frequency given by

$$\omega_0 = \sqrt{\beta / m} \quad \text{...(1)}$$

Here, β is the force constant and m is the mass of the electron.

The displacement x of the electron produced by the application of a field E_{loc} is given by

$$-eE_{loc} = \beta x = m\omega_0^2 x. \quad \text{...(2)}$$

The static electronic polarizability is

$$\alpha_e = \frac{p}{E_{loc}} = -\frac{ex}{E_{loc}} = \frac{e^2}{m\omega_0^2}. \qquad ...(3)$$

The electron polarizability will depend on frequency as shown below.

We treat the system as a simple harmonic oscillator. Let ω_0 be the resonance frequency of the oscillator. Let ω be the frequency of the local electric field.

The equation of motion in the local electric field $E_{loc} \sin \omega t$ is

$$m\frac{d^2x}{dt^2} + m\omega_0^2 x = -eE_{loc} \sin \omega t. \qquad ...(4)$$

Substituting $x = x_0 \sin \omega t$, we obtain

$$m(-\omega^2 + \omega_0^2)\, x_0 = -eE_{loc}. \qquad ...(5)$$

The dipole moment has the amplitude

$$p_0 = -ex_0 = \frac{e^2 E_{loc}}{m(\omega_0^2 - \omega^2)}. \qquad ...(6)$$

The electronic polarizability is

$$\alpha_e = \frac{p_0}{E_{loc}} = \frac{e^2/m}{(\omega_0^2 - \omega^2)}. \qquad ...(7)$$

But in the visible region the frequency dependence (dispersion) is not usually very important in most transparent materials.

In quantum theory the expression corresponding to Eq. (7) is

$$\alpha_e = \frac{e^2}{m}\sum_j \frac{f_{ij}}{(\omega_{ij}^2 - \omega^2)}.$$

Here, f_{ij} is called the *oscillator strength* of the electric dipole transition between the atomic states i and j.

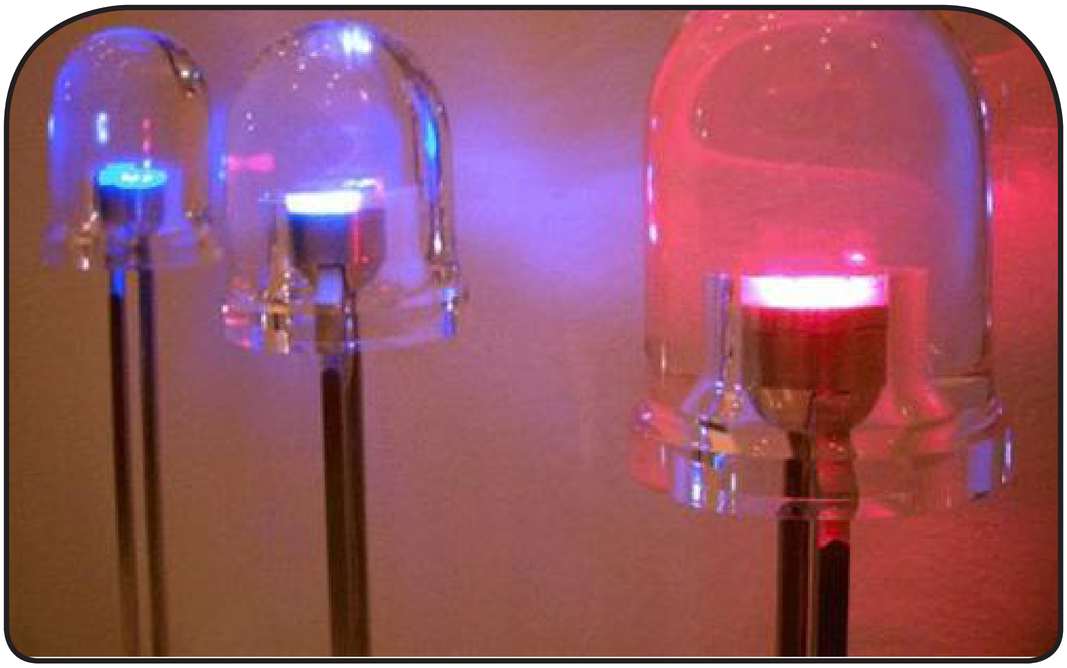
EXERCISE

1. Explain what is meant by polarization in dielectrics.
2. Distinguish the following terms carefully: (1) the applied electric field, (2) the macroscopic electric field in a dielectric, (3) the local field in a dielectric.
3. What is meant by local field in a solid dielectric? Deduce an expression for the local field for structures possessing cubic symmetry.
4. Obtain Clausius – Mossotti formula relating macroscopic dielectric constant with microscopic polarizabilities.
5. Give a schematic sketch of the variation of the total polarizability of an atom as a function of the frequency, explaining the physical origin of the various contributions and the relevant frequency ranges.

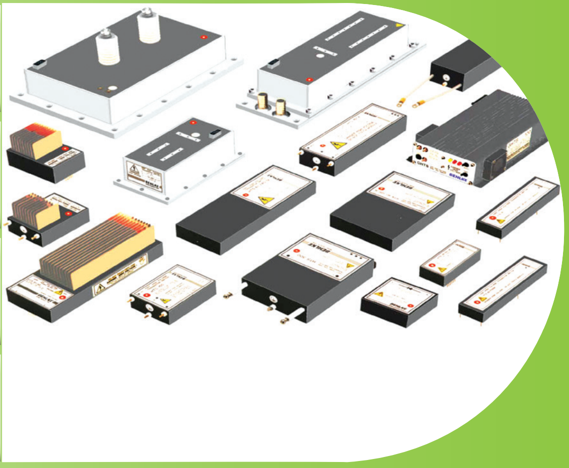
PART – VII

SOLID STATE DEVICES

55. Semiconductors
56. Optoelectronic Devices
57. Hall Effect



CHAPTER 55



SEMICONDUCTORS

AT A GLANCE

55.1 Electron and Hole Concentrations in Intrinsic Semiconductors in Thermal Equilibrium
55.2 Fermi Level in N-type Extrinsic Semiconductor
55.3 Fermi Level in P-type Extrinsic Semiconductor
55.4 Electrical Conductivity
55.5 Diffusion Current
55.6 Total Current
55.7 Carrier Generation and Recombination : Carrier Life Times
55.8 The Continuity Equation
55.9 Derivation of Expressions for the Width of Depletion Layer and Height of Potential Barrier
55.10 Derivation of Voltage Current (V-I) Equation for a P-N Junction Diode
55.11 P-N Diode Clipping Circuits
55.12 P-N Diode Clamping Circuits
55.13 Differentiating Circuit
55.14 Integrating Circuit
55.15 Metal-Semiconductor Junction
55.16 Schottky Diode

55.1 Electron and Hole Concentrations in Intrinsic Semiconductors in Thermal Equilibrium

Consider an intrinsic semiconductor in thermal equilibrium at temperature T.K. We shall now calculate the number of free electrons excited into the conduction band and the hole concentration in the valence band.

To calculate the number of electrons excited into conduction band at temperature T, we assume:

1. The energy is measured from the top of the valence band.
2. The effect of the lattice is simply to modify the free electron mass from m to m_n and m_p for the conduction and valence bands, respectively.
3. The electron density of states in the conduction band is equal to that for free electrons, *i.e.*,

$$N_n(E) = \frac{1}{2\pi^2}\left(\frac{2m_n}{\hbar^2}\right)^{3/2} (E - E_g)^{1/2} \text{ per unit volume.} \qquad \text{...(1)}$$

The Fermi function is

$$f(E) = \frac{1}{1 + e^{(E-E_F)/kT}} \qquad \text{...(2)}$$

The total number of electrons with energy between E and $E + dE$ is given by the product of $N(E)\,dE$ and the Fermi function. So the total number of electrons in the conduction band n is given by

$$n = \int_{E_g}^{\infty} \frac{1}{2\pi^2}\left(\frac{2m_n}{\hbar^2}\right)^{3/2} (E - E_g)^{1/2} f(E)\,dE \qquad \text{...(3)}$$

$E_g - E_F >> kT$. So the Fermi function may, to a good aproximation, be written as,

$$f(E) = \left[1 + e^{(E-E_F)/kT}\right]^{-1} = e^{-(E-E_F)/kT}$$

$$n = \frac{1}{2\pi^2}\left(\frac{2m_n}{\hbar^2}\right)^{3/2} \int_{E_g}^{\infty} (E - E_g)^{1/2} e^{-(E-E_F)/kT} \qquad \text{...(4)}$$

$$\therefore \qquad n = 2\left(\frac{m_n kT}{2\pi\hbar^2}\right)^{3/2} e^{(E_F - E_g)/kT} \qquad \text{...(5)}$$

This relation gives the density or concentration of electrons in the conduction band of an intrinsic semiconductor.

Concentration of holes in the valence band

The density of hole states in the valence band is given by

$$N_p(E) = \frac{1}{2\pi^2}\left(\frac{2m_p}{\hbar^2}\right)^{3/2} (E)^{1/2} \text{ per unit volume} \qquad \text{...(6)}$$

The probability of finding a hole with energy E is just $1 - f(E)$.

$$\begin{aligned} f_p(E) &= 1 - f(E) \\ &= 1 - \frac{1}{1 + e^{(E-E_F)/kT}} \\ &= 1 - \left[1 + e^{(E-E_F)/kT}\right]^{-1} = 1 - \left[1 - e^{(E-E_F)/kT}\right] \\ &= e^{(E-E_F)/kT} \end{aligned}$$

Hence the density of holes in the valence band

$$p = \int_{-\infty}^{0} \frac{1}{2\pi^2}\left(\frac{2m_p}{\hbar^2}\right)^{3/2} (E)^{1/2} e^{(E-E_F)/kT}\,dE \qquad \text{...(7)}$$

$$\therefore \qquad p = 2\left(\frac{m_p kT}{2\pi\hbar^2}\right)^{3/2} e^{-E_F/kT} \qquad ...(8)$$

This is the hole concentration in the valence band.

Law of Mass Action

Multiplying Eqs. (5) and (8), we get

$$np = 4\left(\frac{kT}{2\pi\hbar^2}\right)^3 (m_n m_p)^{3/2} e^{-E_g/kT} = AT^3 e^{-E_g/kT} \qquad ...(9)$$

Here, E_g is the *width of forbidden energy gap* between conduction and valence bands and

$$A = 4\left(\frac{k}{2\pi\hbar^2}\right)^3 (m_n m_p)^{3/2} \rightarrow \text{a constant.}$$

Eq. (9) shows that the *product of hole and electron densities depends on temperature T and forbidden energy gap E_g but is independent of the Fermi level E_F*.

Thus the product of electron and hole concentrations, for a given material, is constant at a given temperature. If an impurity is added to increase n, there will be a corresponding decrease in p such that product np remains constant.

Since for an intrinsic semiconductor, $n = p = n_i$, we arrive at an important relationship, called the *law of mass action*.

$$np = n_i^2 = AT^3 e^{-E_g/kT} \qquad ...(10)$$

Here n_i is called the intrinsic density of either carrier.

Eq. (10) is true for a semiconductor regardless of the donor or acceptor concentrations.

Fermi Level in Intrinsic Semiconductors

For intrinsic semiconductors, $n = p = n_i$

Therefore, from Eqs. (5) and (8) we get

$$2\left(\frac{m_n kT}{2\pi\hbar^2}\right)^{3/2} e^{(E_F - E_g)/kT} = 2\left(\frac{m_p kT}{2\pi\hbar^2}\right)^{3/2} e^{-E_F/kT}$$

$$m_n^{3/2} e^{(E_F - E_g)/kT} = m_p^{3/2} e^{-E_F/kT}$$

or $$E_F = \frac{E_g}{2} + \frac{3}{4} kT \ln\left(\frac{m_p}{m_n}\right) \qquad ...(11)$$

If $m_p = m_n$, then $E_F = E_g/2$.

This shows that the Fermi level E_F lies exactly in the centre of the forbidden energy gap E_g, as shown in Fig. 55.1.

Electrical Conductivity

The electrical conductivity of intrinsic semiconductors in the not very low temperature range is due to intrinsic charge carriers, *i.e.*, to electrons and holes. Such conductivity is sometimes termed as *intrinsic conductivity*.

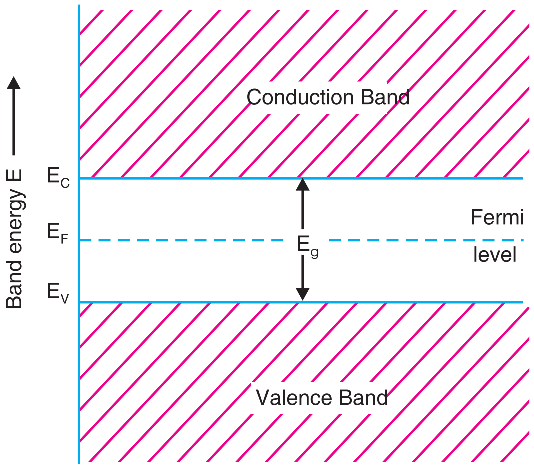


Fig. 55.1

There are two types of carriers in the intrinsic semiconductor. So its specific conductance is the sum of the conductivity $\sigma_n = |e|\, n\, \mu_n$ due to free electrons, with the concentration n and mobility μ_n, and of the conductivity $\sigma_p = |e|\, p\, \mu_p$ due to the presence of holes, with the concentration p and mobility μ_p. The mobility is the magnitude of the drift velocity per unit electric field:

$$\mu = |v|\, E \qquad \text{...(12)}$$

The mobility is defined to be positive for both electrons and holes, although their drift velocities are opposite. In an ideal intrinsic semiconductor the mobility is determined by collisions between electrons and phonons.

The electrical conductivity of an intrinsic semiconductor is

$$\sigma = |e|\,(n\,\mu_n + p\,\mu_p) \qquad \text{...(13)}$$

Since for an intrinsic semiconductor $n = p$, we have

$$\sigma_i = |e|\, n\,(\mu_n + \mu_p) \qquad \text{...(14)}$$

Here σ_i denotes the intrinsic conductivity.

55.2 Fermi Level in N-type Extrinsic Semiconductor

When a small amount of pentavalent impurity is added to the crystal, it creates extra electrons without adding any new holes. This impurity introduces new energy levels into the energy band picture. The location of these new levels is slightly below the bottom of the conduction band for intrinsic semiconductor (the gap was forbidden to electrons of the intrinsic crystal only). The width of forbidden band in germanium is 0.72 eV.

The energy required to move an electron from a donor impurity into the conduction band is of the order of 0.01 eV. Since at normal ambient temperature the thermal energy is considered to be about 0.02 eV, it is concluded that almost all the electrons are detached from donor atoms and have conduction band energies. The energy band diagram for a N-type semiconductor is shown in Fig. 55.2. Here E_D represents the energy level corresponding to donor impurities.

In an intrinsic (pure) semiconductor, the Fermi level E_F lies in the middle of the energy gap between the valence band and the conduction band indicating equal concentration of free electrons and holes. When a donor type impurity is added to the crystal, then if we assume that all the donor atoms are ionized, the donor electrons will occupy the states near the bottom of the conduction band. Hence it will be more difficult for the electrons from the valence band to cross the energy gap by thermal agitation. Consequently, the number of holes of the valence band is decreased. Since Fermi level is a measure of the probability of occupancy of the allowed energy states, E_F for N-type semiconductor must move closer to the conduction band, as shown in Fig. 55.2.

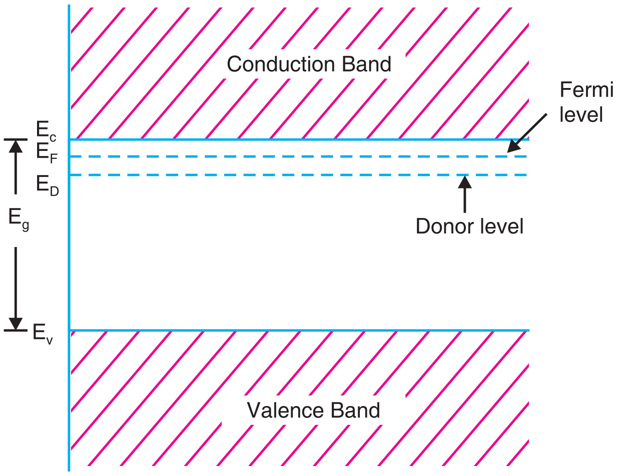


Fig. 55.2

Let the density of ionized donor atoms be N_d. Then at usual temperatures all the donor levels will be fully activated and the donor atoms will be ionized. It means that the density of electrons in the conduction band will be approximately equal to the donor atoms density, *i.e.*, $n \approx N_d$. Therefore, we have

$$n = N_d = 2\left(\frac{m_n kT}{2\pi\hbar^2}\right)^{3/2} e^{(E_F - E_C)/kT} = N_c\, e^{(E_F - E_C)/kT}$$

where $$N_c = 2\left(\frac{m_n\, kT}{2\pi\hbar^2}\right)^{3/2} \rightarrow \text{a constant}$$

$$\therefore \qquad \frac{N_c}{N_d} = e^{-(E_F - E_C)/kT}$$

Taking the logarithm of both the sides,

$$\log_e \frac{N_c}{N_d} = -\frac{(E_F - E_C)}{kT}$$

$$\therefore \qquad E_F = E_C - k\,T \log_e \frac{\quad}{\quad}$$

This shows that the *Fermi level lies below the bottom of the conduction band.*

As temperature rises, the Fermi level goes on falling below E_C.

As temperature is sufficiently raised, the electrons and holes generated due to thermal agitation increase significantly and at a stage become fully dominant over the extrinsic carriers. Then the Fermi level approaches the middle of forbidden energy gap.

55.3 Fermi Level in P-type Extrinsic Semiconductor

When an acceptor-type impurity is added to the crystal, it modifies the energy level diagram of the semiconductor. The presence of impurity creates new energy levels which are in the gap in the neighbourhood of the top of the valence band of energies (Fig. 55.3). Ambient temperature results in ionization of most acceptor atoms and thus an apparent movement of holes from the acceptor levels to the valence band. The energies for holes are highest near the valence band and decrease vertically upward in the energy level diagram. Alternatively, one might say that electrons are accepted by the acceptors, these electron are supplied from the valence band, leaving a preponderance of holes in the valence band.

The energy band diagram for a *p*-type semiconductor is shown in Fig. 55.3. The acceptor level is shown by E_A.

In a *p*-type crystal, the concentration of holes in the valence band is more than the concentration of electrons in the conduction band. Therefore, the Fermi level shifts towards the valence band (Fig. 55.3). The acceptor level lies immediately above the Fermi level.

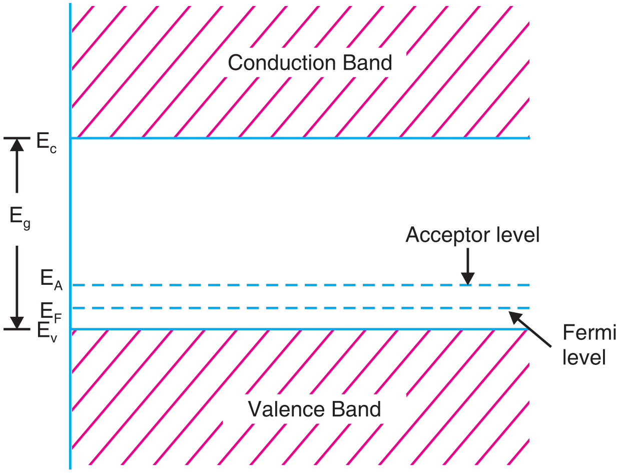


Fig. 55.3

If we assume that only acceptor atoms are present and all are ionized, we have $p \approx N_a$.

So $$p = N_a = 2\left(\frac{m_p kT}{2\pi\hbar^2}\right)^{3/2} e^{(E_v - E_F)/kT}$$

$$\boxed{= N_V\, e^{(E_V - E_F)/kT}}$$

where $$N_V = 2\left(\frac{m_p kT}{2\pi\hbar^2}\right)^{3/2} \rightarrow \text{a constant}$$

$$\therefore \qquad \frac{N_v}{N_a} = e^{-(E_V - E_F)/kT}$$

Taking log, we get

$$\log \frac{}{} = -(E_V - E_F)/kT$$

$$\therefore \qquad E_F = E_V + kT \log_e \frac{N_v}{N_a}.$$

This shows that the *Fermi level lies above the top of valence band*.

The position of Fermi level depends upon temperature and the number of impurity atoms. When the number of impurity atoms increases, the number of holes in the valence band will increase and the Fermi level will shift towards the valence band. When the concentration of acceptor impurity atoms is sufficiently high, the number of holes will be far greater than the conduction electrons and the Fermi level may even lie in the valence band. As the temperature is sufficiently increased, electrons from the valence band are excited to the conduction band and finally the p-type crystal will start behaving like an intrinsic semiconductor when the number of electrons in the conduction band will be nearly equal to the valence holes. Thus at extremely *high temperatures the Fermi level will shift towards the middle of forbidden energy gap*.

55.4 Electrical Conductivity

Electrical conductivity of an intrinsic semiconductor is

$$\sigma = |e|\,(n\mu_n + p\mu_p).$$

For an intrinsic semiconductor, $n = p = n_i$.

Therefore, *conductivity of intrinsic semiconductor*

$$\sigma_{int} = e\, n_i\,(\mu_n + \mu_p).$$

In an N-type semiconductor, concentration of minority charge carriers, *i.e.*, holes is negligible in comparison with the concentration of majority charge carriers, *i.e.*, electrons. Thus $n >> p$ and, therefore, the *conductivity of n-type semiconductor*

$$\sigma_{n-type} \approx en\mu_n \approx eN_d\,\mu_n$$

because $n = N_d \rightarrow$ density of ionized donor atoms.

For a p-type semiconductor, $p >> n$ and its conductivity is given by

$$\sigma_{p-type} \approx ep\mu_p \approx eN_a\,\mu_p$$

because $p = N_a \rightarrow$ density of ionized acceptor atoms.

55.5 Diffusion Current

It is possible to have a nonuniform concentration of particles (electrons and holes) in a semiconductor. In such a case, particles tend to move from a region of higher concentration to a region of lower concentration. This movement of particles due to concentration gradient is called *diffusion*. The current so constituted is called *diffusion current*.

Let us consider that concentration p of holes varies with distance x in the semiconductor in such a way that the hole concentration is greater for lower values of x. Then there is a concentration gradient – dp/dx in the density of holes resulting in a hole flow along the positive x-direction. It represents a *diffusion current*.

The diffusion hole current density J_p is proportional to the concentration gradient and is given by

$$J_p = -eD_p \frac{dp}{dx} \qquad ...(1)$$

where e is electronic charge and D_p is called the *diffusion constant* or *diffusivity for holes*. The negative sign in Eq. (1) appears because $\frac{dp}{dx}$ is negative and J_p is positive along positive x-direction.

A similar equation can be written for diffusion electron current density J_n by replacing p and D_p by n and D_n respectively and by replacing minus sign by a plus sign. Thus

$$J_n = -e D_n \frac{dn}{dx}. \qquad ...(2)$$

55.6 Total Current

In semiconductors, when an external electric field E is applied, both the drift and diffusion currents may occur simultaneously. The total hole and electron currents can then be written as the sum of drift current and diffusion current. Then total current densities for holes and electrons are given by

$$J_p = e\mu_p pE - eD_p \frac{dp}{dx}$$

and

$$J_n = e\mu_n nE - eD_n \frac{dn}{dx}.$$

55.7 Carrier Generation and Recombination: Carrier Life Times

In an intrinsic semiconductor, the number of holes is equal to the number of free electrons. Electron-hole pairs are being produced continuously by thermal energy, say at a rate of g electron-hole pairs per unit volume per second. Simultaneously, other hole-electron pairs disappear as a result of recombination, *i.e.*, free electrons falling into empty covalent bonds. The time period for which an electron (or hole) exists before recombination is called the *mean life time*. On an average, a hole exists for τ_p seconds known as the mean life time of hole. An electron exists for τ_n seconds known as *mean life time* of electron.

Consider a specimen of N-type semiconductor. Let the thermal equilibrium concentrations of holes and electrons be p_0 and n_0 respectively.

Assume that at $t = t'$, the specimen is illuminated (Fig. 55.4).

Now additional hole-electron pairs are generated uniformly throughout the crystal. When the equilibrium is reached, let the new concentrations be $\overline{p}$ and $\overline{n}$. So, the excess concentration of holes is $\overline{p} - p_0$ and excess of concentration of electrons is $\overline{n} - n_0$. Since the radiation causes hole-electron pairs to be created,

increase in hole concentration = increase in electron concentration

$$\overline{p} - p_0 = \overline{n} - n_0 \qquad ...(1)$$

When the steady state is reached (let at time $t = 0$), the illumination is suddenly stopped.

Increase in hole concentration per second due to thermal generation = g, ...(2)

Decrease in hole concentration per second due to recombination = p/τ_p, ...(3)

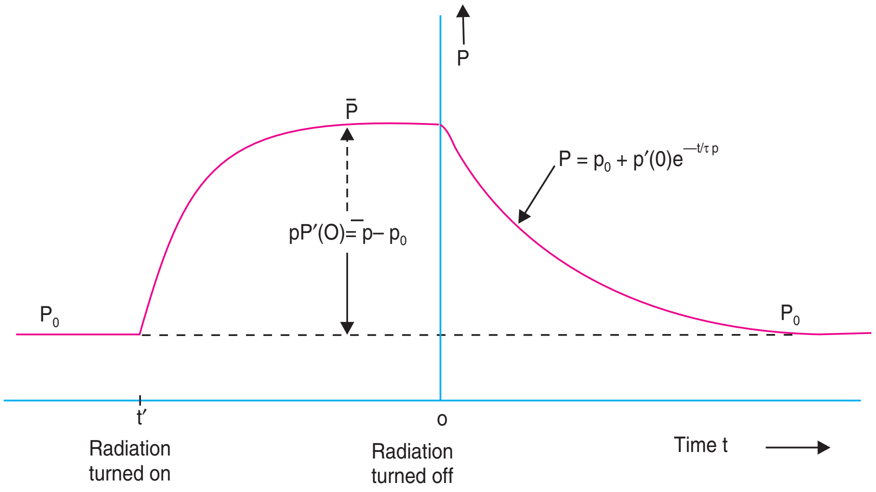


Fig. 55.4

The time rate of increase of hole concentration per second

$$\frac{dp}{dt} = g - \frac{p}{\tau_p} \qquad \text{...(4)}$$

Under steady state condition $dp/dt = 0$ and with no radiation incident on the specimen, the hole concentration p reaches the thermal equilibrium value p_0, *i.e.*,

$$0 = g - \frac{p_0}{\tau_p}$$

or
$$g = \frac{p_0}{\tau_p} \qquad \text{...(5)}$$

Eq. (4) becomes,

$$\frac{dp}{dt} = \frac{p_0}{\tau_p} - \frac{p}{\tau_p} = \frac{p_0 - p}{\tau_p} \qquad \text{...(6)}$$

The excess carrier density p' is defined as the increase in minority concentration above the equilibrium value. Since p' is a function of time, we have

$$p' = p - p_0 = p'(t) \qquad \text{...(7)}$$

or
$$(dp'/dt) = (dp/dt) \qquad (\because\ p_0 \text{ is constant}) \qquad \text{...(8)}$$

Combining Eqs. (6), (7) and (8), we get

$$\frac{dp'}{dt} = \frac{p_0 - p}{\tau_p} = -\frac{p'}{\tau_p} \qquad \text{...(9)}$$

Eq. (9) shows that the *rate of change of excess concentration is proportional to the excess concentration itself.*

The significance of the negative sign is that the change is a decrease in case of recombination and increase in case the concentration is recovering from a sudden depletion.

In our case, the radiation has caused an initial excess concentration $p'(0) = \left(\bar{p} - p_0\right)$ at time $t \le 0$ and subsequently at $t = 0$ the excitation is suddenly removed. Hence the solution of Eq. (9) is given by

$$p'(t) = p'(0)e^{-t/\tau_p} = (\bar{p} - p_0)e^{-t/\tau_p} = p - p_0 \qquad \text{... (10)}$$

The excess concentration decreases exponentially to zero ($p' = 0$ or $p = p_0$) with a time constant τ_p as shown in Fig. 55.4.

Thus, the *excess carrier concentration decreases to zero exponentially with time*.

55.8 The Continuity Equation

The continuity equation states a condition of dynamic equilibrium for the density of mobile carriers in any elementary volume of the semiconductor. We have seen that on disturbing the equilibrium concentrations of carriers in a semiconductor, the concentrations of holes or electrons vary with time approaching the equilibrium value exponentially. In general, the carrier concentration in the body of a semiconductor is a function of both time and distance. The differential equation governing this functional relationship is called continuity equation. This equation is based on the fact that charge can neither be created nor destroyed.

Derivation

Consider an infinitesimal element of volume of unit area and length dx (Fig. 55.5). It is possible to create charge carriers in excess to the equilibrium value by optical excitation (by shining light on the material). Let mobile charge carriers be produced in it at a rate g at some time. Simultaneously, they disappear at a rate r through the process of recombination when free electrons fall into empty covalent bonds. The particles may also flow away from a point at a rate which depends upon the differential current dJ/dx.

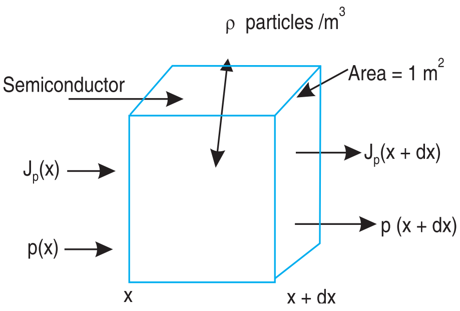


Fig. 55.5

Adding all these particle-current components, we get the time rate of change of particle density ρ.

$$\frac{d\rho}{dt} = g - r - \frac{dJ}{dx} \qquad \text{...(1)}$$

But diffusion current J is given by

$$J = D\frac{d\rho}{dx} \qquad \text{...(2)}$$

Here D is called the *diffusion constant* or *diffusivity*.

Therefore, combining Eqs. (1) and (2), we get the ***diffusion equation***,

$$\frac{d\rho}{dt} = g - r + D\frac{d^2\rho}{dx^2} \qquad \text{...(3)}$$

If in addition to diffusion, drift also takes place simultaneously, then J will represent the total electric current and given by

$$J = -D\frac{d\rho}{dx} + \mu E\rho \qquad \text{...(4)}$$

Here μ is the mobility (average drift velocity per unit electric field) of a carrier.

Then Eq. (1) is modified to

$$\frac{d\rho}{dt} = g - r + D\frac{d^2\rho}{dx^2} - \mu\frac{d(\rho E)}{dx} \qquad \text{...(5)}$$

If no field is applied to the semiconductor and it is in thermal equilibrium with its surroundings, then under steady state, $J = 0$, $\dfrac{d\rho}{dt} = 0$ and space charge density ρ reaches the thermal equilibrium value ρ_0.

Then Eq. (1) gives

$$g - r = 0 \qquad \text{or} \qquad g = r.$$

But rate of recombination r is proportional to charge density, *i.e.*,

$$r = k\rho_0$$

Then in thermal equilibrium, the rate of generation of mobile carriers is given by

$$g = r = k\rho_0 = \frac{\rho_0}{\tau_p} \qquad \text{[See Eq. (5) of Article 55.7]}$$

Here, τ_p is the mean life time of the charge carriers.

Let due to some reason, say by illuminating the specimen by light, the space charge density increase to ρ from its thermal equilibrium value ρ_0. Then

$$g - r = \frac{\rho_0}{\tau_p} - \frac{\rho_0 + \delta\rho}{\tau_p} = -\frac{\delta\rho}{\tau_\rho} = -\frac{\rho - \rho_0}{\tau_\rho} \qquad ...(6)$$

Combining Eqs. (5) and (6), we get

$$\frac{d\rho}{dt} = -\frac{\rho - \rho_0}{\tau_p} + D\frac{d^2\rho}{dx^2} - \mu\frac{d(\rho E)}{dx} \qquad ...(7)$$

Eq. (7) is known as *continuity equation* for charge.

The continuity equation for electrons in N-type semiconductor is obtained by replacing ρ by –*en* and is given by

$$\frac{dn}{dt} = -\frac{n - n_0}{\tau_n} + D\frac{d^2 n}{dx^2} - \mu\frac{d(nE)}{dx} \qquad ...(8)$$

Here, n is the average electron concentration within the volume.

Similarly, continuity equation for holes is obtained by replacing p by ep and is given by

$$\frac{dp}{dt} = -\frac{p - p_0}{\tau_p} + D\frac{d^2 p}{dx^2} - \mu\frac{d(pE)}{dx} \qquad ...(9)$$

Here, p is the average hole concentration within the volume.

55.9 Derivation of Expressions for the Width of Depletion Layer and Height of Potential Barrier

Fig. 55.6 shows a *P-N* junction.

Fig. 55.6 (a) shows *P*-type and *N*-type semiconductors before they are joined. *P*-type semiconductor has negative acceptor ions (shown by encircled minus sign) and positively charged free holes which move about on *P*-side. Similarly, *N*-type semiconductor has positive donor ions (shown by encircled positive sign) and negatively charged free electrons which move about *N* side.

Fig. 55.6 (b) shows the *P-N* junction after the two pieces are joined.

P-type material has a high concentration of holes. *N*-type material has high concentration of free electrons. Hence there is a tendency of holes to diffuse over to *N*-side and electrons to *P*-side. The process is known as *diffusion*. So due to diffusion, some of the holes from *P*-side cross over to *N*-side where they combine with electrons and become neutral. Similarly, some of the electrons from *N*-side cross over to *P*-side where they combine with holes and become neutral. Thus a region is formed which is known as *depletion layer* or *space charge region* because there is no charge available for conduction. The thickness of the depletion layer is usually of the order of 10^{-6} m.

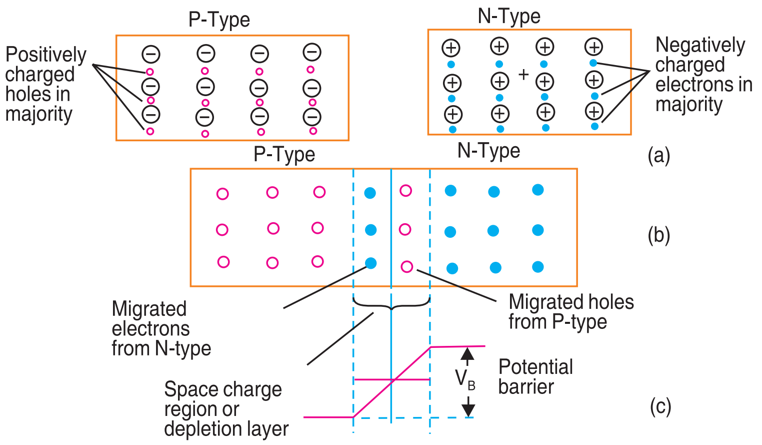

P-Type
N-Type
Positively charged holes in majority
Negatively charged electrons in majority
(a)
P-Type
N-Type
(b)
Migrated electrons from N-type
Migrated holes from P-type
Space charge region or depletion layer
V_B
Potential barrier
(c)

Fig. 55.6

The diffusion of holes and electrons continues till a potential barrier [Fig. 55.6 (c)] is developed in space charge region which prevents further diffusion or neutralization.

Junction Voltage. When the depletion layer is formed, there are negative immobile ions in *P*-type semiconductor and positive immobile ions in *N*-type semiconductor (Fig. 55.7).

Now due to charge separation, a voltage V_B is developed across the junction under equilibrium condition. This voltage is called *junction voltage* or *internal voltage*. The potential barrier V_B set up in this manner gives rise to electric field. This electric field prevents the respective majority carriers from crossing the barrier region. The potential barrier is about 0.3 V for Ge and 0.7 V for Si at room temperature (300 K).

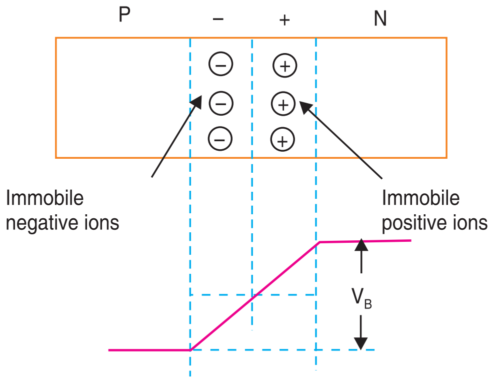

P
–
+
N
Immobile negative ions
Immobile positive ions
V_B

Fig. 55.7

As soon as the *P-N* junction is formed, the diffusion of holes (from *P* region to *N* region) and electrons (from *N* region to *P* region) takes place for a short time. Then, a force (called a potential barrier) is automatically developed across the junction. The *potential barrier* stops the further diffusion of holes and electrons from one side to the other. The difference of potential from one side of the barrier to the other side is called the *height of the barrier*.

Fig. 55.8 shows the variation of potential with distance along the junction.

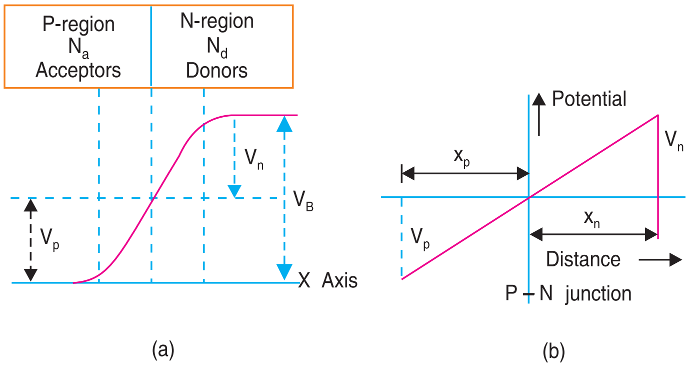

P-region
N_a
Acceptors
N-region
N_d
Donors
V_n
V_B
V_p
X Axis
(a)
Potential
x_p
V_n
V_p
x_n
Distance
P N junction
(b)

Fig. 55.8

The barrier potential V_B is given by

$$V_B = |V_P| + |V_n| \qquad \text{...(1)}$$

Here, V_p = magnitude of the potential fall in P region,

V_n = magnitude of the potential rise in N region.

The width of the depletion region is given by

$$x = x_p + x_n \qquad \text{...(2)}$$

Here, x_p and x_n are the widths of depletion layer in P and N regions respectively.

To find the distribution of barrier potential in the space charge region, we use Poisson's equation

$$\frac{d^2V}{dx^2} = -\frac{\rho}{\varepsilon} \qquad \text{...(3)}$$

Here ρ is the volume density of charge and ε is the permittivity of the medium.

The charge density in P-side of depletion layer is given by

$$\rho = -eN_a.$$

Here, N_a is the density of completely ionized acceptor atoms.

Substituting the value of ρ in Eq. (3), we get

$$\frac{d^2V}{dx^2} = \frac{eN_a}{\varepsilon}$$

Integrating it, we get

$$\frac{dV}{dx} = \frac{eN_a . x}{\varepsilon} + C_1 \qquad \text{...(4)}$$

Here, C_1 is a constant of integration.

Applying boundary condition :

at $x = -x_p, \frac{dV}{dx} = 0$

we get $C_1 = \frac{eN_a}{\varepsilon}, x_p$

Substituting the value of C_1 in Eq. (4), we get

$$\frac{dV}{dx} = \frac{eN_a}{\varepsilon}.x + \frac{eN_a}{\varepsilon}.x_p$$

or

$$\frac{dV}{dx} = \frac{eN_a}{\varepsilon}. (x + x_p) \qquad \text{...(5)}$$

Integrating Eq. (5), we get

$$V = \frac{eN_a}{\varepsilon}.\left(\frac{x^2}{2} + x_p x\right) + C_2 \qquad \text{...(6)}$$

Here, C_2 is another constant of integration.

Its value is obtained by applying the condition that at $x = 0$, $V = 0$.

or $C_2 = 0.$

$$\therefore \quad V = \frac{eN_a}{\varepsilon}.\left(\frac{x^2}{2} + x_p x\right) \qquad \text{...(7)}$$

For P-region, at $x = -x_p$, $V = V_p$.

$$V_P = \frac{eN_a}{\varepsilon}.\left(\frac{x_p^2}{2} - x_p^2\right) = -\frac{eN_a\, x_p^2}{2\varepsilon}$$

$$\therefore \qquad |V_p| = \frac{eN_a}{2\varepsilon}.x_p^2 \qquad ...(8)$$

Similarly, the Poisson’s equation for *N*-side of depletion layer is

$$\frac{d^2V}{dx^2} = -\frac{\rho}{\varepsilon} = \frac{eN_d}{\varepsilon}.$$

Here, N_d is the density of completely ionized donor atoms.

Proceeding as above and applying boundary conditions :

at $x = 0, V = 0$

and at $x = x_n, dV/dx = 0$ and $V = V_n$

we get $V_n = -\frac{eN_d}{2\varepsilon} x_n^2$

$$|V_n| = \frac{eN_d}{2\varepsilon} x_n^2 \qquad ...(9)$$

Therefore, the height of potential barrier is given by

$$V_B = |V_p| + |V_n|$$

$$= \frac{eN_a}{2\varepsilon} x_p^2 + \frac{eN_d}{2\varepsilon} x_n^2$$

$$\therefore \qquad V_B = \frac{e}{2\varepsilon}\left[N_a x_p^2 + N_d x_n^2\right] \qquad ...(10)$$

Since the crystal as a whole is electrically neutral, the number of charge carriers on both sides must be equal, *i.e.*,

$$N_a x_p = N_d x_n$$

or $x_n = \frac{N_a}{N_d} . x_p$

Substituting the value of x_n in Eq. (10), we get

$$V_B = \text{—}\left[N_a x_p^2 + N_d \text{——} \times x_p^2\right]$$

or $V_B = \frac{e}{2\varepsilon} N_a . x_p^2 \left[1 + \frac{N_a}{N_d}\right]$ ...(11)

or $x_p = \left[\frac{2\varepsilon V_B}{eN_a\{1 + (N_a / N_d)\}}\right]^{1/2}$

$$\therefore \qquad x_p = \left[\frac{2\varepsilon V_B}{e} . \frac{N_d / N_a}{N_a + N_d}\right]^{1/2} \qquad ...(12)$$

Similarly, $x_n = \left[\frac{2\varepsilon V_B}{eN_d\{1 + (N_d / N_a)\}}\right]^{1/2}$

$$\therefore \qquad x_n = \left[\frac{2\varepsilon V_B}{e}\cdot\left(\frac{N_a/N_d}{N_a+N_d}\right)\right]^{1/2} \qquad \text{... (13)}$$

Therefore, total width of depletion layer

$$x = x_p + x_n$$

$$x = \left\{\frac{2\in V_B}{e(N_a+N_d)}\right\}^{1/2}\left[\left(\frac{N_d}{N_a}\right)^{1/2}+\left(\frac{N_a}{N_d}\right)^{1/2}\right] \qquad \text{... (14)}$$

If $N_d >> N_a$, then $(N_a + N_d) \rightarrow N_d$ and $N_a/N_d \rightarrow 0$.

Eq. (14) becomes

$$x = \left(\frac{2\varepsilon V_B}{eN_d}\right)^{1/2}\left(\frac{N_d}{N_a}\right)^{1/2} = \left(\frac{2\varepsilon V_B}{eN_a}\right)^{1/2} \qquad \text{...(15)}$$

Eq. (15) gives the width of depletion layer.

The width of depletion layer decreases with increase in impurity concentration.

Expression for the height of barrier potential in terms of impurity densities

The density of electrons in conduction band in an intrinsic semiconductor is given by

$$n = n_c = 2\left(\frac{2\pi m_e kT}{h^2}\right)^{3/2} e^{(E_F-E_c)/kT} \qquad \text{...(16)}$$

Addition of donor impurity raises the Fermi level by ΔE on *N*-side.

The new Fermi level is

$$E'_F = E_F + \Delta E$$

Expression for density of electrons on *N*-side modifies to

$$(n_c)_N = 2\left(\frac{2\pi m_e kT}{h^2}\right)^{3/2} e^{(E_F+\Delta E-E_c)/kT}$$

$$\text{or} \qquad (n_c)_N = ne^{\Delta E/kT} \qquad \text{...(17)}$$

Similarly, the density of electrons on *P*-side is given by

$$(n_c)_P = ne^{(-\Delta E'/kT)} \qquad \text{...(18)}$$

because addition of donor impurities brings down the Fermi level on *P*-side by $\Delta E'$.

From Eqs. (17) and (18), we get

$$\frac{(n_c)_N}{(n_c)_P} = ne^{(\Delta E+\Delta E')/kT}$$

$$= e^{(eV_B/kT)} \qquad \text{...(19)}$$

Here, V_B is the height of potential barrier and $eV_B = \Delta E + \Delta E'$.

From Eq. (19), we get

$$V_B = \frac{kT}{e}\log_e\left[\frac{(n_c)_N}{(n_c)_P}\right] \qquad \text{...(20)}$$

By the *law of mass action*,

$$n_c n_h = np = n_i^{\,2}.$$

Here n_i is called the intrinsic density of either carrier.

For a semiconductor having N_d completely ionised donor atoms, it becomes

$$(n_c' + N_d)\, n_h' = n_i^2.$$

Here, n_c' and n_h' are the number of thermally generated electron-holes.

For an N-type semiconductor, $N_d >> n_c'$ and we have

$$N_d N_h' = n_i^2.$$

Therefore, the hole concentration in N-type semiconductor, *i.e.*, minority carrier concentration

$$(n_h')_N = \frac{n_i^2}{N_d} \qquad ...(21)$$

Similarly, for a semiconductor having N_a completely ionised acceptor atoms,

$$n_c' (\text{Na} + n_h') = n_i^2.$$

But for a P-type semiconductor $N_a >> n_h'$ and we have

$$n_c' N_a = n_i^2.$$

Therefore, electron concentration in P-type material

$$(n_c')_P = \frac{n_i^2}{N_a} \qquad ...(22)$$

If we assume that minority carriers on P-side are thermally generated, then

$$(n_c')_P = \frac{(n_c)_P}{N_a} \qquad ...(23)$$

Eq.(20) for V_B becomes

$$V_B = \frac{kT}{e} \log_e \left[\frac{(n_c)_N}{n_i^2 / N_a} \right]$$

or

$$V_B = \frac{kT}{e} \log_e \left[\frac{(n_c)_N N_a}{n_i^2} \right] \qquad ...(24)$$

Since majority carriers (electrons) in N-type semiconductor are largely contributed by donor atoms, we take

$$(n_c)_N = N_d$$

Therefore, Eq. (24) becomes

$$V_B = \frac{kT}{e} \log_e \left[\frac{N_d N_a}{n_i^2} \right] \qquad ...(25)$$

This gives the height of *barrier potential in terms of impurity densities that create it.*

55.10 Derivation of Voltage Current (V-I) Equation for a P-N Junction Diode

The density or the concentration of electrons in the N-region is

$$n_n = 2\left(\frac{2\pi m_n^* kT}{h^2} \right)^{3/2} e^{-(E_{cn} - E_{fn})/kT} \qquad ...(1)$$

Here, m_n^* = effective mass of the electron in the conduction band,

E_{cn} = energy at the bottom of conduction band,

and $\boxed{E_{fn} = \text{Fermi level.}}$

The *P*-region also contains some thermally generated electrons of concentration

$$n_p = 2\left(\frac{2\pi m_n^* kT}{h^2}\right)^{3/2} e^{-(E_{cp}-E_{fp})/kT} \qquad \text{...(2)}$$

$$\therefore \qquad \frac{n_n}{n_p} = \frac{e^{-(E_{cn}-E_{fn})/kT}}{e^{-(E_{cp}-E_{fp})/kT}} \qquad \text{...(3)}$$

For the same semiconducting material, the positions of the conduction and valence band edges do not change with doping. But the position of the Fermi level changes with both concentration and type of doping. Thus for *N* and *P* regions,

$$E_{cn} = E_{cp} \text{ and } E_{fn} \neq E_{fp}.$$

Therefore, Eq. (3) reduces to

$$\frac{n_n}{n_p} = \mathrm{e}^{(E_{fn} - E_{fp})/kT} = e^{(eV_B/kT)} \qquad \text{...(4)}$$

Here, $E_{fn} - E_{fp} = eV_\mathrm{B} = E_B =$ barrier energy.

$V_B =$ barrier potential.

Thus at equilibrium, the minority carrier electron concentration in the *P*-region is related to the majority carrier electron concentration in the *N*-region by

$$n_p = n_n e^{(-eV_B/kT)} \qquad \text{...(5)}$$

On applying a small forward voltage *V*, the effective barrier potential decreases to $(V_B - V)$ and some of the majority carriers diffuse across the junction. Thus holes diffuse to *N*-side and electrons to *P*-side. Consequently, minority carrier electron concentration at the edge of transition region on the *P*-side becomes greater than its equilibrium value by a small amount Δn_p. The majority carrier electron concentration in the *N*-region, however, remains nearly constant.

Thus under the forward bias condition, Eq. (5) becomes

$$\begin{aligned} n_p + \Delta n_p &= n_n\, e^{-e(V_B - V)/kT} \\ &= n_p\, e^{e(V_B)/kT} \,.\, e^{-e(V_B - V)/kT} \qquad \text{[using Eq. (5)]} \\ &= n_p\, e^{(eV/kT)} \end{aligned}$$

or

$$\Delta n_p = n_p\,[e^{eV/kT} - 1)] \qquad \text{...(6)}$$

Similarly, for excess holes on the *N*-side

$$\Delta p_n = p_n\,[e^{eV/kT} - 1] \qquad \text{...(7)}$$

The diffusion currents produced by these charges can be determined by using diffusion equations and Fick's first law.

The electron diffusion current injected into the *P*-region at the junction is given by

$$I_n = \frac{eAD_n}{L_n}\Delta n_p = \frac{eAD_n}{L_n} \,.\, n_p\left[e^{eV/kT} - 1\right] \qquad \text{...(8)}$$

Here, $A =$ area of cross-section of the junction,

$D_n =$ diffusion coefficient of electron,

$L_n =$ electron diffusion length.

Similarly, the hole diffusion current injected into the *N*-region at the junction is

$$I_p = \frac{eAD_p}{L_p}\, p_n\left[e^{eV/kT} - 1\right] \qquad \text{...(9)}$$

Here, D_p and L_p are respectively the diffusion coefficient and diffusion length for holes.

Since the direction of both the currents is the same, the total diode current

$$I = I_P + I_n = eA\left[\frac{D_P}{L_p}p_n + \frac{D_n}{L_n}n_p\right]\left[e^{eV/kT} - 1\right]$$

$$\therefore \qquad I = I_0(e^{eV/kT} - 1) \qquad \text{...(10)}$$

Here, $$I_0 = eA\left[\frac{D_P}{L_P}P_n + \frac{D_n}{L_n}n_p\right] \qquad \text{...(11)}$$

Eq.(10) is called the *diode equation*.

I_0 = *reverse saturation current* at temperature T K.

The empirical form of Eq. (10) is

$$I = I_0[e^{eV/\eta kT} - 1] \qquad \text{...(12)}$$

Here, η is a numerical constant which depends upon the material of the diode.

For *Ge*, $\eta = 1$. For *Si*, $\eta = 2$.

(*i*) For large forward bias voltage, $e^{Ve/\eta kT}$ is large when compared to 1.

$$\therefore \qquad I_f = I_0 e^{Ve/\eta kT} \qquad \text{...(13)}$$

Thus *forward current I_f increases exponentially with voltage V* except for very small values of V (Fig. 55.9).

(*ii*) For reverse bias voltage, V is negative and Eq. (12) becomes

$$I_R = I_0\left[\frac{1}{e^{Ve/\eta kT}} - 1\right] \qquad \text{...(14)}$$

For large reverse bias voltage, exponential term quickly becomes negligible with respect to unity. Then

$$I_R \approx -I_0 \qquad \text{...(15)}$$

Thus *reverse current I_R is constant independent of applied reverse bias*.

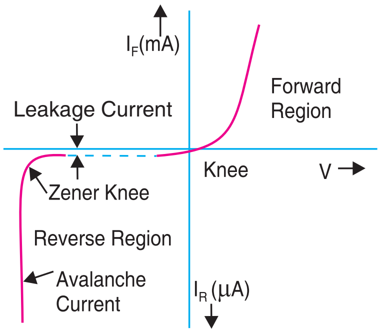


Fig. 55.9

WAVE SHAPING CIRCUITS

55.11 P-N Diode Clipping Circuits

A wave shaping circuit which controls the shape of output waveform by removing or clipping a portion of the applied wave is known as clipping circuit.

"A clipping circuit is one in which a sine wave form is shaped by removing (clipping) a portion of the input sine wave".

Clippers are used in radar, digital and other electronic devices.

Clipping Circuits are used in Radar.

1. (*i*) Positive Clipper :

In a positive clipper, positive half cycle of the input voltage is removed. Fig. 55.10 shows the circuit arrangement of a positive clipper.

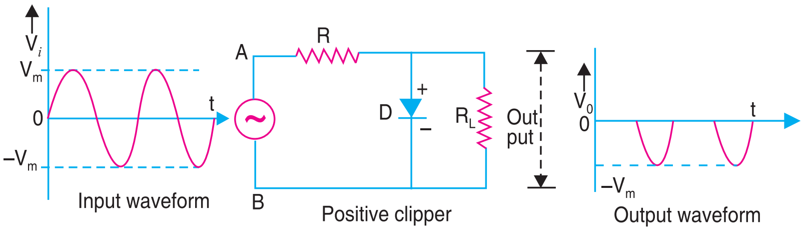


Fig. 55.10

(*i*) ***Working:*** When the positive half cycle of the signal voltage is applied to the clipper, *i.e*, when *A* is positive with respect to *B*, the diode *D* is forward-biased. Hence it acts like a closed switch (or short) and conducts heavily. Therefore the voltage drop across the diode or across the load resistor R_L is zero. So, the output voltage during the positive half-cycles is zero, *i.e*., positive half cycle is clipped off (removed).

(*ii*) In the negative half cycle of the signal voltage, *i.e*., when *B* is positive with respect to *A*, the diode *D* is reverse-biased. Hence it acts like an open switch. Consequently, the entire input voltage appears across the diode or across the load resistor R_L. So current flows through the load R_L. Hence there is an output voltage.

(ii) Negative Clipper.

In a negative clipper, negative half cycle of the input voltage is removed.

Fig. 55.11 shows the circuit of a negative clipper.

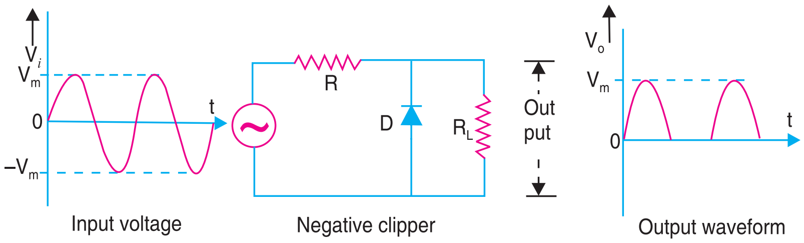


Fig. 55.11

***Working* :**

(*i*) During the positive half cycle of the input voltage, the diode is reverse-biased. It is not conducting. So current flows through the load R_L. Hence there is an output voltage.

(*ii*) During the negative half cyle, the diode is forward-biased and so it is conducting heavily. Hence no current flows through the load. There is no output voltage, *i.e*., the negative half of the input voltage is removed.

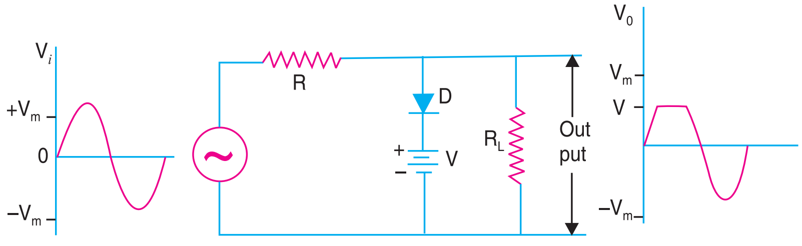
2. Biased Clipper

A biased clipper is used when it is desired to remove a small portion of positive or negative half cycle of the signal voltage.

(*i*) Biased Positive Clipper

It clips off a small portion of the positive half cycle of the input voltage at a low voltage level *V*.

Fig. 55.12 shows the circuit of a biased positive clipper.



Input voltage Biased positive clipper Output waveform

Fig. 55.12

Working :

(*i*) *When the input signal voltage is less than V for positive half cycle*: In this case, the diode *D* is reverse-biased, *i.e.*, acts as an open circuit. Therefore, most of the input voltage appears across the output.

(*ii*) *When the input signal voltage is greater than V for positive half cycle* : In this case, the diode *D* is forward-biased, *i.e.*, conducts heavily. The output voltage is equal to +*V*. The output voltage will stay at +V as long as the input voltage is greater than +*V*.

(*iii*) *When the input signal voltage is negative*: In this case, the diode *D* is reverse-biased. Hence entire input voltage appears across the output.

(*ii*) Biased Negative Clipper

It removes any desired portion of the negative half cycle of the input voltage.

Fig. 55.13 shows the circuit of a biased negative clipper.

The circuit is same as that for the positive clipper, except that the polarities of the diode and the associated bias supply are reversed.

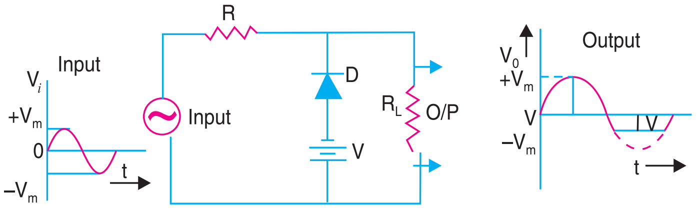


Fig. 55.13

Working :

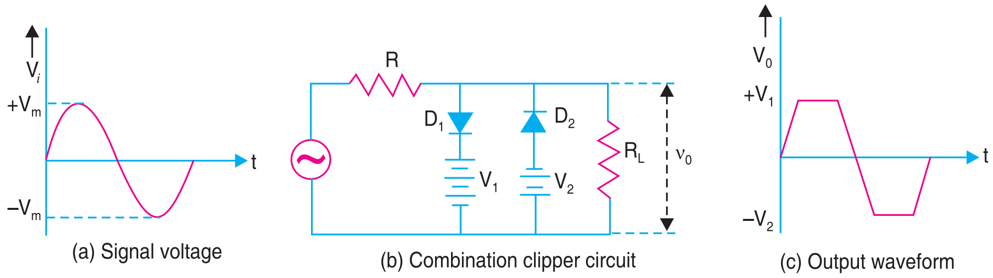
During the entire positive half cycle, the diode *D* is reverse-biased and so it does not conduct. Hence the entire input voltage appears across the load.

During the negative half cycle, as long as the input voltage is less than *V*, the diode is reverse-biased. So, current flows through R_L and there is an output voltage. When the input voltage exceeds *V*, *D* becomes forward-biased and so it conducts. Hence output voltage remains at *V* equal to bias battery voltage.

(3) Biased Combination Clipper

A combined circuit of biased positive and negative clippers is called combination clipper. This is shown in Fig. 55.14.

With the combination clipper, a portion of both positive and negative half cycles of input voltage can be removed.



(a) Signal voltage (b) Combination clipper circuit (c) Output waveform

Fig. 55.14

Working :

(*i*) When the input voltage lies between $+V_1$ and $-V_2$, neither of the diodes conducts and the input voltage appears across the load, without any change.

(*ii*) When positive input voltage is greater than $+V_1$, diode D_1 is forward-biased while diode D_2 is reverse-biased. Therefore, a voltage $+V_1$ appears across the load and positive signal voltage above $+V_1$ is clipped off.

(*iii*) During negative half cycle, the diode D_2 is forward-biased and D_1 is reverse-biased so long as the input negative voltage is greater than $-V_2$. So the output stays at $-V_2$. Hence, signal voltage beyond $-V_2$ volt is clipped off.

55.12 P-N Diode Clamping Circuits

Clamping is a process of introducing a dc level into an ac signal as shown in Fig. 55.15. *A clamping circuit is a device that places either the positive or negative peak of a signal at a desired level*.

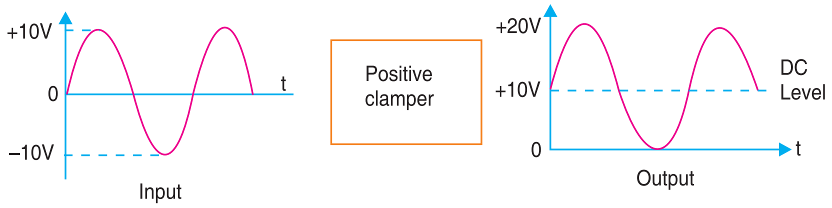
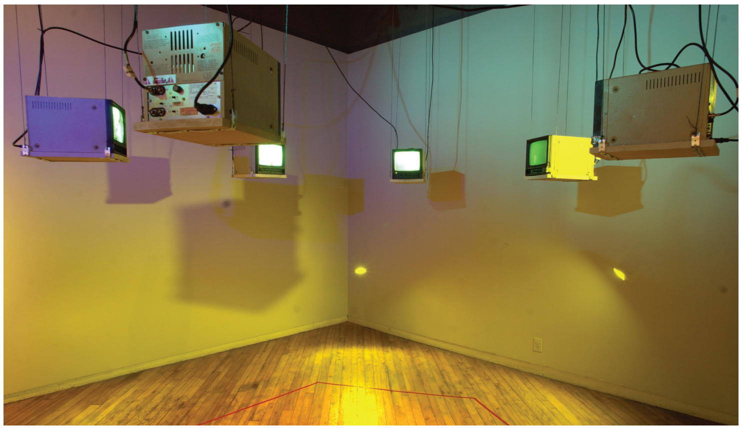


Fig. 55.15

The waveforms of input and output remain the same. But there is a vertical shift either upward or downward.

"A circuit that holds either the positive or negative peak of a signal at a desired dc level is known as clamping circuit".

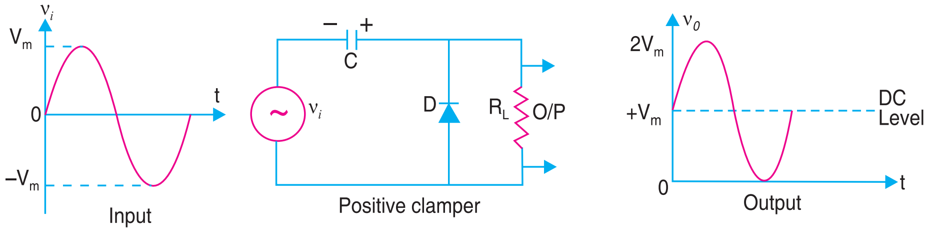
Clamping circuits are widely used in TV receivers to add a dc voltage to the video signal.



Clamping Circuits are Used in TV receivers.

(1) Positive Clamper

When a clamper shifts the original signal in vertical upward direction, it is known as positive clamper. Fig. 55.16 shows the circuit of a positive clamper. It consists of a diode *D* and a capacitor *C*. The clamped output is taken across the load R_L.



Input

Positive clamper

Output

Fig. 55.16

Here we assume that diode is ideal, *i.e.*, it exhibits an arbitrarily sharp break at 0 volt and that its forward resistance is zero. The input signal is sinusoidal which begins at $t = 0$. The capacitor is uncharged at $t = 0$. Here our aim is to find the waveform of the output v_0.

During the negative half cycle of the input voltage, the diode *D* is forward-biased and conducts heavily. At the maximum negative peak of applied signal, the capacitor is charged to V_m with polarities as shown in Fig. 55.16.

During the positive half cycle, the diode *D* is reverse-biased and does not conduct. The capacitor charged to V_m behaves as a battery.

Hence the output voltage v_0 is given by

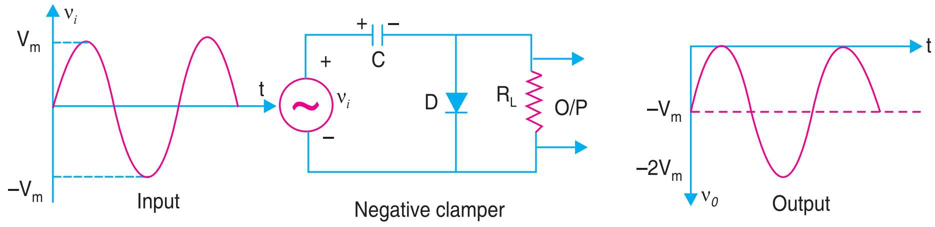
$$v_0 = \text{output voltage} = V_m + v_i.$$

According to the changes in v_i, there would be corresponding changes in output voltage.

It is obvious that $V_m + v_i$ simply shifts the input voltage upward by V_m. The output voltage is obtained by shifting the input voltage upwards by V_m. The output is thus positively clamped voltage.

(2) Negative Clamper

When a clamper shifts the original signal in vertical downward direction, it is known as negative clamper. The input wave is pushed (*i.e.*, shifted) downwards, so that the positive peak lies on the zero level. Fig. 55.17 shows the circuit of negative clamper.



Input

Negative clamper

Output

Fig. 55.17

Working :

During the positive half cycle, diode *D* is forward-biased and conducts heavily. At the maximum positive peak of the applied signal, the capacitor is charged to V_m with polarities as shown in Fig. 55.17.

The voltage across the capacitor opposes the input voltage v_i. So the output voltage is given by

$$v_0 = \text{output voltage} = v_i - V_m.$$

This is equivalent to subtracting a constant voltage V_m from every point on the sine curve of v_i over the full cycle.

DIFFERENTIATING AND INTEGRATING CIRCUITS

55.13 Differentiating Circuit

A circuit in which the output voltage is proportional to the derivative of input voltage is known as differentiating circuit.

A differentiating circuit is a simple *RC* series circuit in which the output is taken across the resistor *R* as shown in Fig. 55.18.

The output voltage is directly proportional to the rate of change of the signal input, *i.e.*,

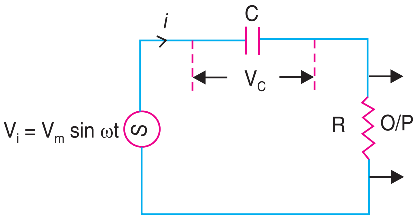


Fig. 55.18

$$\text{output voltage} \propto \frac{d}{dt}\text{(input voltage)}$$

$$V_{out} \propto \frac{dV_{in}}{dt} \quad \text{or} \quad V_{out} = RC\frac{dV_{in}}{dt}$$

The output will be zero for a constant input signal.

(*i*) *RC time constant*. The time constant *RC* must be many times less than the time period of the input, *i.e.*, $RC << T$. For good differentiation, the time constant *RC* of the circuit should be about one-tenth of the interval of the pulses to be differentiated.

(*ii*) The reactance offered by the capacitor (X_c) must be large.

The *AC* input voltage can be represented by

$$V_i = V_m \sin \omega t \qquad \text{...(1)}$$

At any instant, let *i* be the current in the circuit.

Applying Kirchoff's law to the *RC* circuit,

$$V_i = \text{Voltage across } R + \text{voltage across C.}$$

$$V_i = iR + \frac{1}{C}\int_0^t i\,dt \qquad \left(\because V_C = \frac{q}{C} = \frac{1}{C}\int_C^t i\,dt\right)$$

$$\frac{V_i}{R} = i + \frac{1}{RC}\int_0^t i\,dt \qquad \text{...(2)}$$

Since $RC << T$, the second term is much greater than the first term. So first term can be neglected.

$$\therefore \quad \frac{V_i}{R} = \frac{1}{RC}\int_0^t i\,dt \qquad \text{...(3)}$$

Differentiating Eq. (3) w.r.t. time, we get

$$\frac{1}{R}\frac{d}{dt}(V_i) = \frac{1}{RC}i$$

$$RC\frac{d}{dt}(V_i) = iR = V_0 \qquad (\because iR = V_0)$$

$$\text{or} \quad V_0 = RC\frac{d}{dt}(V_i) \qquad \text{...(4)}$$

$$\therefore \qquad V_0 \propto \frac{d}{dt}(V_i) \qquad (\because RC \text{ is constant})$$

Thus the output voltage is directly proportional to the derivative of the input voltage.

(*i*) The output will be zero for a constant input signal.

(*ii*) If the amplitude of the input changes abruptly, the differentiated wave is a sharp narrow pulse.

(*iii*) For the less abrupt change in input, the output is a pulse approaching a very narrow rectangular form.

(*iv*) For the constant rate of change in the input, the output has a constant value.

(*v*) The waveform of the output depends upon the time constant *RC* of the circuit, and the frequency and the shape of the input voltage.

Output Waveforms

(*i*) The input signal with a triangular waveform [Fig. 55.19 (a)] will produce a square wave [Fig. 55.19 (b)] output.

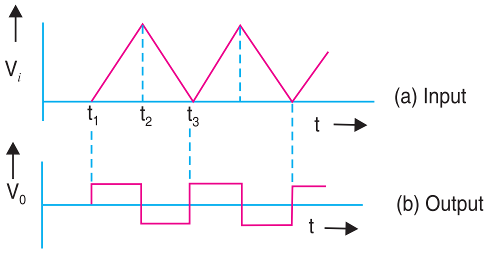


Fig. 55.19

(*ii*) The square wave input signal will produce a series of equally spaced pulses called spikes [Fig. 55.20].

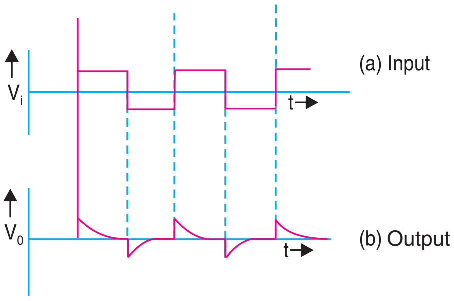


Fig. 55.20

(*iii*) The input signal with a sawtooth waveform will produce a rectangular wave output [Fig. 55.21].

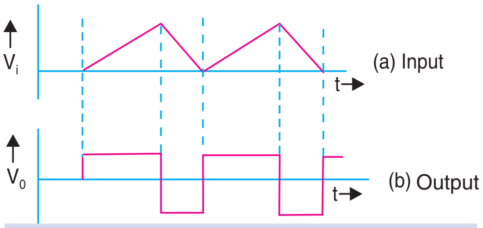


Fig. 55.21

(*iv*) The rectangular wave input signal will produce a series of spikes with alternately long and short spacing [Fig. 55.22].

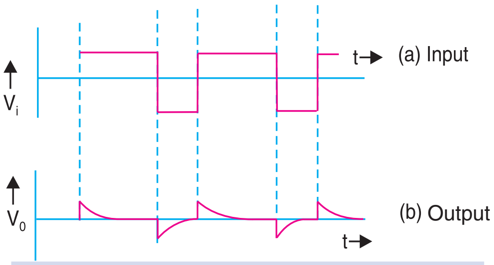


Fig. 55.22

55.14 Integrating Circuit

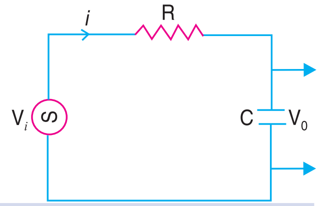


Fig. 55.23

A circuit in which the output voltage is directly proportional to the integral of the input voltage is known as integrating circuit, *i.e.,*

output voltage $\propto \int$ input voltage .

An integrating circuit is a simple *RC* circuit in which the output is taken across the capacitor *C*, as shown in Fig. 55.23.

The output voltage is directly proportional to the time integral of the input signal, *i.e.*,

$$V_{out} \propto \int V_{in}\, dt ,$$

or $$V_{out} = \frac{1}{RC}\int V_{in}\, dt \;.$$

The waveform of the output will depend upon the time constant and frequency and shape of the input signal. For good integration, the time constant *RC* of the circuit should be about ten times the interval of the pulses to be integrated. When the time constant of *RC* circuit is long, each succeeding pulse adds to the charge on the capacitor until it reaches an average value of the applied voltage. When the time constant is short, the capacitor charges and discharges so rapidly that the output waveform is nearly the same as the input waveform with the exception of the edges being rounded off.

DALLAS MAXIM
MAX5945

Integrating Circuit

The *AC* input voltage can be represented by ... (1)

$$V_i = V_m \sin \omega t$$

At any instant, let *i* be the current in the circuit.

Applying Kirchoff's law to the *RC* circuit.

$$V_i = \text{voltage across } R + \text{ voltage across } C = iR + \frac{1}{C}\int_0^t i\, dt$$

or $$CV_i = RCi + \int_0^t i\, dt \qquad \text{...(2)}$$

Since $RC >> T$ (Time Period), the second term can be neglected.

$\therefore$ $$CV_i = RCi \qquad \text{...(3)}$$

Integrating Eq. (3) w.r.t. *t*, we get

$$C\int V_i\, dt = RC\int_0^t i\, dt \qquad \text{...(4)}$$

But output voltage V_0 = voltage across capacitor $C = \frac{1}{C}\int_0^t i\, dt$

or $$\int_0^t i\, dt = C.V_0 \qquad \text{...(5)}$$

From Eqs. (4) and (5),

$$C\int V_i\,dt = RC.CV_0$$

or
$$V_0 = \frac{1}{RC}\int V_i\,dt$$

$\therefore$
$$V_0 \propto \int V_i dt \qquad \text{...(6)} \qquad (\because RC \text{ is constant})$$

Output Waveforms

(*i*) A square wave input will produce a triangular wave output (Fig. 55.24).

(*ii*) An input signal with equally spaced pulses will produce a square wave.

(*iii*) The rectangular wave input signal will produce a sawtooth wave output.

(*iv*) An input signal with un-equally spaced pulses will produce a rectangular wave output.

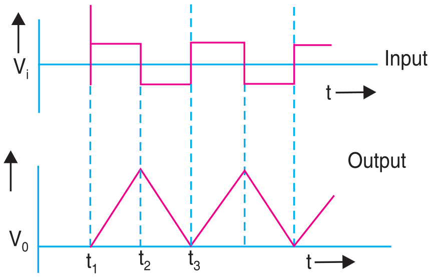


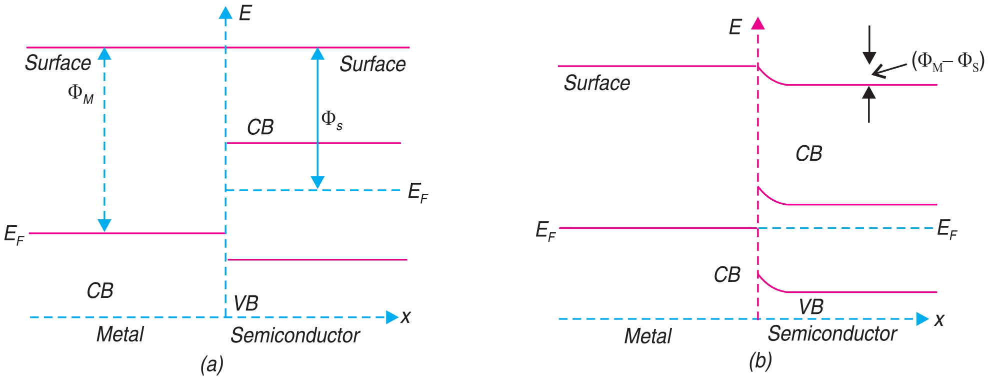
Fig. 55.24

55.15 Metal-Semiconductor Junction

Consider a junction between a metal and a semiconductor. When two materials come into contact, the energies at their surface are initially equal. The energy difference between the Fermi level and the surface is called the *work function* ϕ. If an electron is at the Fermi level, the work function is the energy required to completely remove the electron from the material.

Consider a metal with work function ϕ_M in contact with an *n*-type semiconductor with work function ϕ_s (Fig. 55.25). Let $\phi_M > \phi_s$.

At initial contact, energies at surfaces are equal [Fig. 55.25 (a)]



(a) (b)

Fig. 55.25

The Fermi level of the semiconductor is higher than the Fermi level of the metal. Electrons will flow from the *CB* of the semiconductor to the metal, and the metal will become negatively charged. [Fig. 55.25(b)]. At equilibrium, Fermi levels are equal. Now, in order for electrons to pass between the *CB* of the semiconductor and the metal, they must go over a potential barrier. This situation is very much like that encountered in a *p-n* junction.

Forward-biased junction. If we apply a positive voltage V_a to the metal side of the junction, the barrier presented to electrons flowing from the semiconductor into the metal is decreased, and the junction is "forward-biased" [Fig. 55.26(a)].

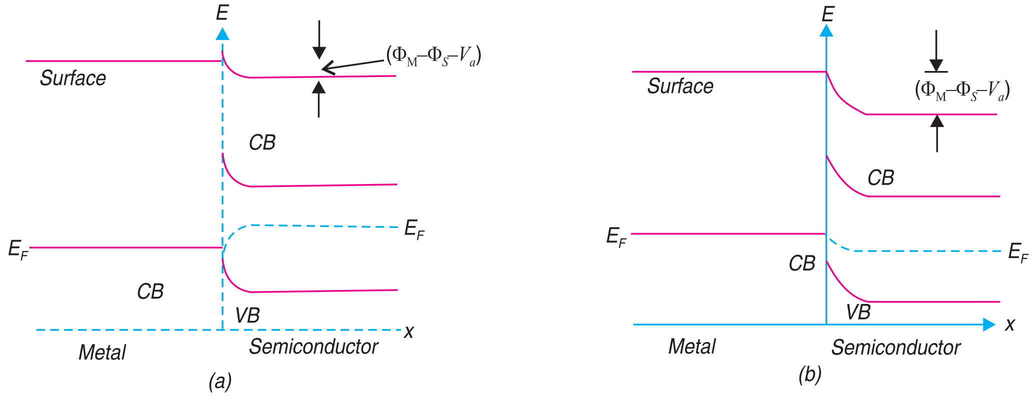


Fig. 55.26

Reverse-biased junction. If we apply a positive voltage to the semiconductor side of the junction (V_a is negative), the barrier presented to electrons flowing from the metal into the semiconductor is unchanged, and very little current flows. The junction is "reverse-biased" [Fig. 55.26(b)]. This metal-semiconductor junction is a *rectifying contact*. Such a contact is used in a device called a *Schottky-barrier* diode.

An aluminium silicon (*Al-Si*) junction forms a rectifying contact. The work function of *Al* is $\phi_M = 4.25$ eV. The work function of *n*-type *Si* is $\phi_s = 3.5eV$.

55.16 Schottky Diode

A metal-semiconductor junction diode is known as **Schottky Diode**.

Construction

It consists of a junction between a metal (like platinum, gold, silver, tungsten etc.) and *N*-type doped semi conductor. The most common semiconductor material used is silicon. GaAs can also be used which has lower noise and higher operating frequency. Schottky diode has no depletion layer. Fig. 55.27 shows the diode and its schematic symbol.

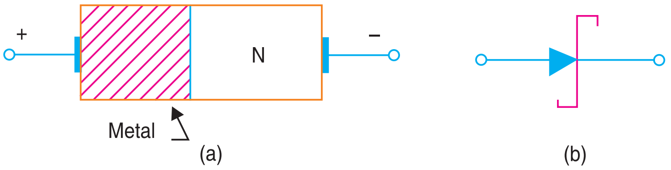
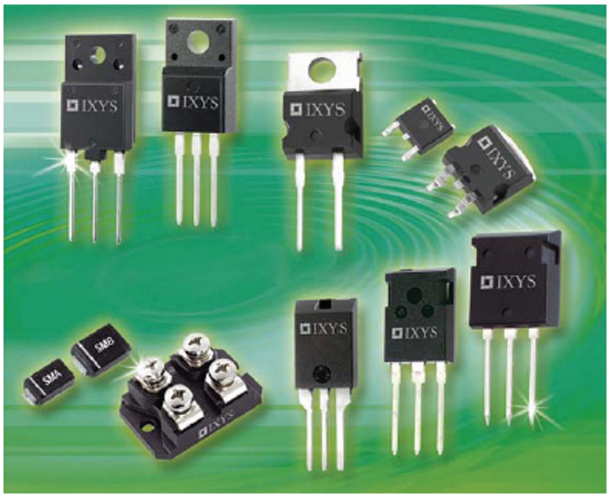


Fig. 55.27

Operation

Metals have a work function which is defined as the minimum energy required for an electron to escape into vacuum. It is the energy difference between the Fermi and vacuum levels. For semiconductors, the energy differences between vacuum level and the bottom of conduction band is known as electron affinity. When a junction is formed between a metal and a semiconductor, the Fermi levels on both sides get aligned and a barrier (to electron flow) is formed due to energy difference between work function of the metal and electron affinity of the semiconductor.



Schottky Diodes

When the diode is unbiased, electrons on N-side have lower energy levels than electrons in the metal. Hence the electrons on the semiconductor side cannot surmount the junction barrier (Schottky barrier) for going over to metal.

When the diode is forward-biased, the electrons on the N-side gain enough energy to cross the barrier and enter the metal. Since the electrons plunge into the metals with very large energy, they are called 'hot-carriers'. Hence the diode is called ***hot carrier diode.***

If a reverse voltage is applied to the diode, the electrons in the metal cannot overcome the barrier. Therefore, no conduction takes place. The metal-semiconductor junction is thus unidirectional and behaves as a diode.

Characteristics

The volt-ampere (V-I) characteristics of Schottky diode is similar to the P-N junction diode when forward-biased. Since there is no charge storage in the metal-semiconductor junction, the diode would switch off as soon as the reverse potential is applied.

Advantages over P-N Junction diodes

Schottky diode has the following advantages over an ordinary P-N junction diode:

(*i*) An ordinary P-N junction diode is a bipolar device because it has both electrons and holes as majority carriers. A Schottky diode has electrons as majority carriers on both sides of the junction. The electrons are majority charge carriers in N-type semiconductor and remain so in the metal after crossing over to the other side. Thus the Schottky diode is a *unipolar* device because there are no minority carriers in the reverse direction (*i.e.*, no significant current from metal to semiconductor when the diode is reverse-biased). The delay due to the hole-electron recombination present in junction diodes is thus absent.

(*ii*) The junction contact area between semiconductor and metal is larger than in point contact diode and hence the forward resistance is lower (*i.e.*, noise is comparatively lower). Low noise figure is highly important in communication receivers and radar units etc.

(*iii*) Since no holes are available in the metal, there is no depletion layer or charge storage. The delay due to the hole-electron recombination present in junction diodes is thus absent here. So Schottky diode can switch OFF faster than a bipolar diode.

Applications

(*i*) Schottky diode can easily rectify signals of frequencies exceeding 300 MHz. The output will be a perfect sine wave as shown in Fig. 55.28.

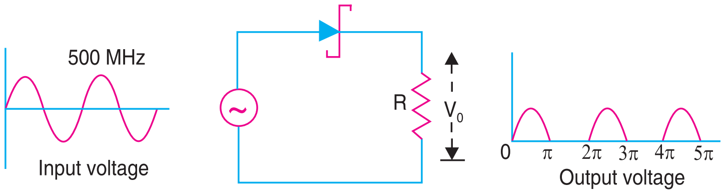


Fig. 55.28

(*ii*) It finds application in very high frequency and fast digital circuits.

(*iii*) It is used in microwave detection and mixing.

(*iv*) It is used in clipping and clamping circuits.

EXERCISE

1. What do you mean by intrinsic semiconductors? Obtain an expression for the intrinsic carrier concentration in an intrinsic semiconductor. Under what condition will Fermi level be in the middle of the forbidden gap ? (*Rajasthan,* 84)

2. What do you mean by intrinsic and extrinsic semiconductors? Show that the Fermi level lies at the centre of the forbidden gap for intrinsic semiconductor. Is the location of Fermi level temperature dependent? (*Rohilkhand,* 1991)

3. What is Fermi level ? Prove that :

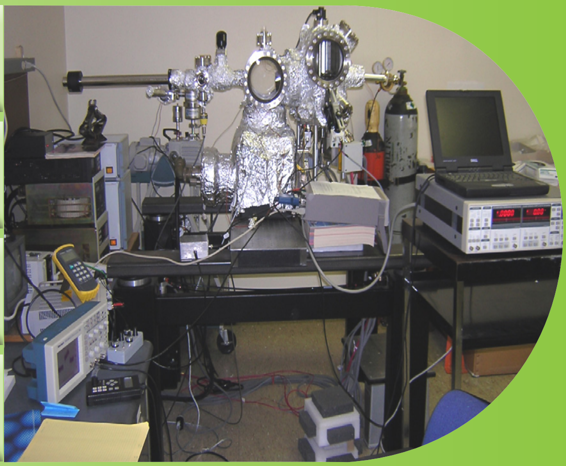
(*i*) Fermi level in an intrinsic semiconductor lies exactly in the centre of the forbidden band.

(*ii*) Fermi level in *N*-type semiconductor lies below the bottom of conduction band.

(*iii*) Fermi level in *P*-type semiconductor lies above the top of valence band.

4. Derive an expression relating the shift in the Fermi level and the carrier density in an extrinsic semiconductor.

5. What do you understand by drift and diffusion currents in the case of a semiconductor?



CHAPTER
56

OPTOELECTRONIC DEVICES

AT A GLANCE

56.1 Photo Diode
56.2 Photo Transistor
56.3 Photoconductors
56.4 Light Emitting Diode (LED)
56.5 Liquid Crystal Display (LCD)
56.6 The Solar Cell

56.1 Photo Diode

Principle. A reverse-biased *P-N* junction diode has a reverse saturation current which is mainly due to the flow of the minority carriers. When light is incident on the depletion region of the reverse-biased *p-n* junction, the concentration of minority carriers (electrons in *p* type and holes in *n* type) increases to a great extent. But the change in majority carriers is too low. Consequently reverse current increases. The reverse current through the diode varies almost linearly with the intensity of light.

A photo diode is essentially a reverse-biased P-N junction diode which is designed to respond to photon absorption.

Construction. A photo diode consists of a *P-N* junction embedded in a clear plastic capsule [Fig. 56.1(*a*)]. The symbol of a photo diode is shown in Fig. 56.1(*b*). Light is allowed to fall upon one surface across the junction. All the sides of the plastic capsule, excepting the illuminated one, are either painted black or enclosed in a metallic case.

Working and Characteristics. When photo diode is kept under dark condition and a sufficient reverse voltage is applied, then an almost constant current, independent of magnitude of reverse bias, is obtained. This current corresponds to the reverse saturation current due to thermally generated minority carriers. It is called *dark current.* It is proportional to the concentrations of minority carriers and is denoted by I_d. Majority charge carriers are not allowed to cross the junction by the potential hill under this reverse bias condition.

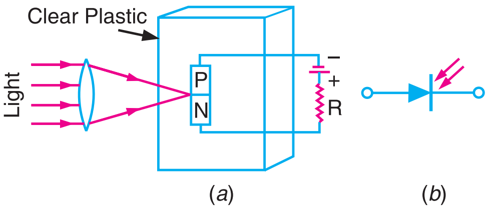


Fig. 56.1

When light falls on the diode surface, additional electron-hole pairs are formed. But since the concentration of majority carriers is much greater as compared to that of minority carriers, the percentage increase of majority carriers is much smaller than the percentage increase of minority carriers. Hence, we can neglect the increase in majority carrier density and can consider the radiation entirely as a *minority carrier injector.* These injected minority carriers diffuse to the junction, cross it and contribute to the additional current.

Thus under large reverse bias conditions, the total reverse current is given by

$$I = I_s + I_d$$

where I_s is the short circuit current and is proportional to light intensity.

With any bias V, the reverse current due to thermal electron-hole pairs, *i.e.*, dark current is, given by

$$I_d = I_0 (1 - e^{Ve/\eta kT}).$$

Hence, the volt-ampere characteristic of photo diode is given by

$$I = I_s + I_0 (1 - e^{Ve/\eta kT})$$

where η is equal to 1 for Ge and equals 2 for Si.

The volt-ampere characteristic curve of a photo diode is as shown in Fig. 56.2. From the curve it is seen that

(*i*) the current increases with increase in the level of illumination for a given reverse voltage.

(*ii*) only for the dark current at zero voltage the current is zero.

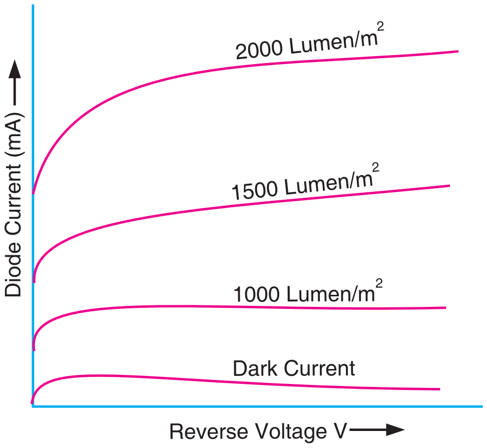


Fig. 56.2

The photo diode finds extensive application in light detection systems, reading of film sound track, light operated switches, high-speed reading of computer punched cards and tapes.

56.2 Photo Transistor

Photo transistor is an *n-p-n* or *p-n-p* transistor in which the forward current is controlled by the incident light at the *pn* junction. Due to transistor action, the current is amplified. For the normal operation of the photo-transistor (*n-p-n*), the base is kept open [Fig. 56.3]. The emitter region is given a negative potential while collector region is given a positive potential. Radiation is concentrated on the region near the collector junction J_c.

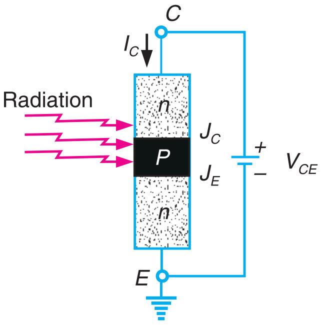


Fig. 56.3

The photo-transistor is usually connected in a common-emitter configuration with the base open. The junction J_E is slightly forward-biased. The junction J_c is reverse-biased. Assume, first, that there is no radiant excitation. Under these circumstances minority carriers are generated thermally. The electrons crossing from the base to the collector, as well as the holes crossing from the collector to the base, constitute the reverse saturation collector current I_{co}. The collector current is given by

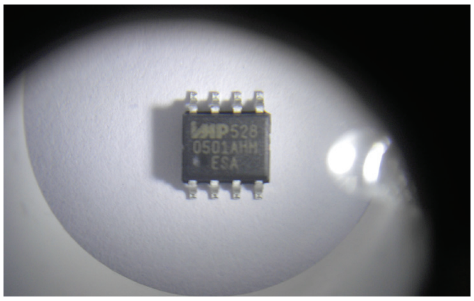


Photo Transistor.

$$I_C = (\beta + 1)\, I_{CO} \quad \text{...(1)}$$

If the light is now turned on, additional minority carriers are photogenerated, and these contribute to the reverse saturation current in exactly the same manner as do the thermally generated minority charges. Let I_L be the component of the reverse saturation current due to the light. Then, the total collector current is

$$I_C = (\beta + 1)\,(I_{CO} + I_L) \quad \text{...(2)}$$

Due to transistor action, the current caused by the radiation is multiplied by the large factor $(\beta + 1)$.

The output volt-ampere characteristics of an *n-p-n* silicon phototransistor are shown in Fig. 56.4 for three different values of illumination intensities.

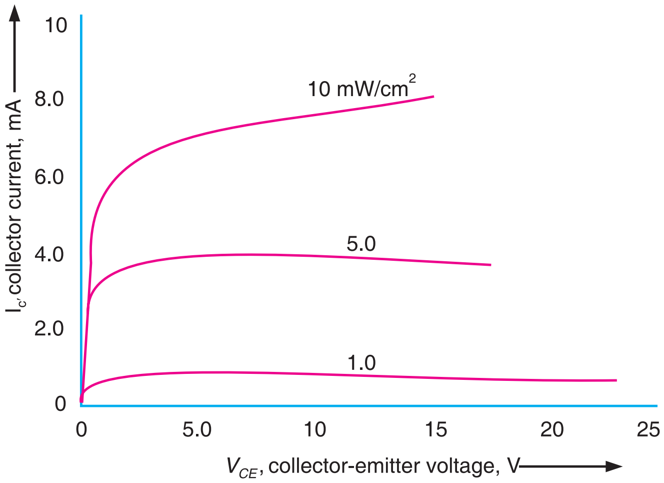


Fig. 56.4

Photo transistors are used in light operated relays. They are also used to compare the intensities of different light sources and for infrared detection.

56.3 Photoconductors

If radiation falls upon a semiconductor, its conductivity increases. This photoconductive effect is explained as follows: Radiant energy supplied to the semiconductor ionizes covalent bonds; that is, these bonds are broken, and hole-electron pairs in excess of those generated thermally are created. These increased current carriers decrease the resistance of the material. Hence such a device is called a *photoresistor*, *or photoconductor.*

Photoconductors

Fig. 56.5 shows the energy diagram of a semiconductor having both acceptor and donor impurities. If photons of sufficient energies illuminate this specimen, *photogeneration* takes place. The following three transitions are possible :

(*i*) An electron-hole pair can be created by a high-energy photon. This is called *intrinsic excitation.*

(*ii*) A photon may excite a donor electron into the conduction band.

(*iii*) A valence electron may go into an acceptor state.

The last two transitions are called *impurity excitations.*

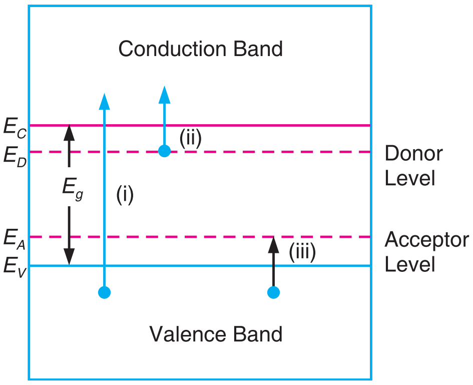


Fig. 56.5

The density of states in the conduction and valence bands is much greater than the density of impurity states. So photoconductivity is due principally to intrinsic excitation.

Cut-off Wavelength and Spectral Response

The minimum energy of a photon required for intrinsic excitation is the forbidden-gap energy (E_g) of the semiconductor material. Hence the minimum frequency ν_c to cause photoconduction is given by

$$h\,\nu_c = E_g$$

Corresponding maximum wavelength or critical wavelength is given by

$$\frac{hc}{\lambda_c} = E_g\,.$$

If λ_c is expressed in microns and E_g in electron volts,

$$\lambda_c = \frac{1{\cdot}24}{E_g}\,.$$

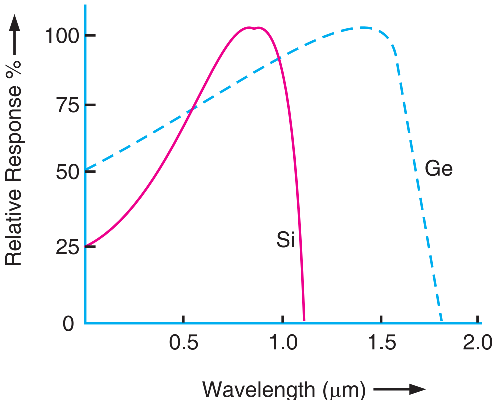
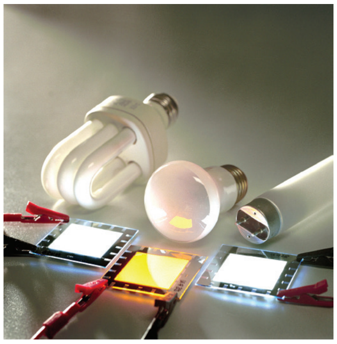


Fig. 56.6

If the wavelength λ of the radiation exceeds λ_c, then the energy of the photon is less than E_g. Such a photon cannot cause a valence electron to enter the conduction band. Hence λ_c is called the *critical*, or *cutoff wavelength*, or *long wavelength threshold*, of the material. For Si, E_g = 1.1 eV and λ_c = 1.13 μm, whereas for Ge, E_g = 0.72 eV and λ_c = 1.73 μm at room temperature.

Fig. 56.6 gives *spectral response* versus wavelength curve for Ge and Si.



Light Emitting Diode

56.4 Light Emitting Diode (LED)

A light emitting diode (LED) is a specially made forward-biased P-N junction diode which emits visible light when energized.

Theory

When a junction diode is forward-biased, electrons from *n*–side and holes from *p*–side move towards the depletion region and they recombine. During this process, energy is released because electrons make transition from conduction band (higher energy level) to valence band (lower energy level) [Fig. 56.7].

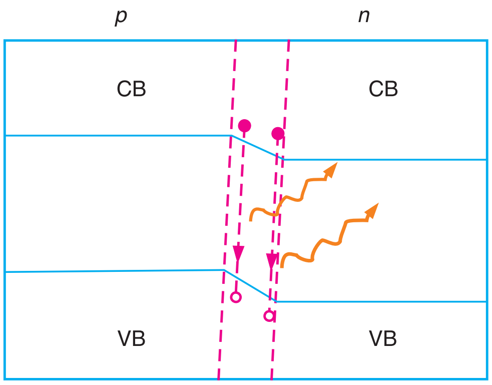


Fig. 56.7

If E_g is the semiconductor band gap, then the energy $E_g = h\nu = \frac{hc}{\lambda}$ is emitted in the form of radiation. The corresponding emission wavelength is given by

$$\lambda = \frac{hc}{E_g}.$$

In case of Gallium Arsenide Phosphide (Ga As P), band gap $E_g = 1.9$ eV, and we get

$$\lambda = \frac{hc}{E_g} = \frac{6 \cdot 62 \times 10^{-34} \times 3 \times 10^{8}}{1 \cdot 9 \times 1 \cdot 6 \times 10^{-19}} \text{ m} = 653 \cdot 3 \,\text{nm} \,(\text{red}) .$$

Construction

At first an *N*-type layer is grown on a substrate and then a *P*-type layer is deposited on it by the process of diffusion [Fig 56.8]. Metal contacts (Anode) are made at the outer edge of the *P*-layer so that more upper surface is left free for light to escape. For making Cathode connections, a metal film (preferably gold) is coated at the bottom of the substrate. This film also reflects as much light as possible to the surface of the device.

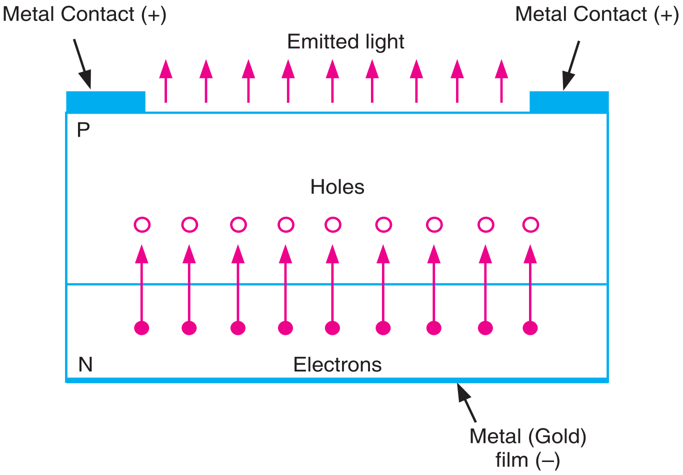
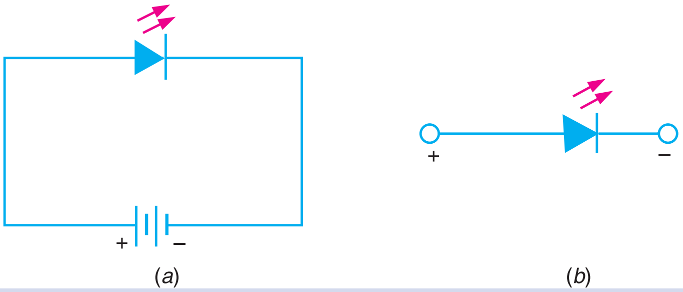


Fig. 56.8

Ga As *P* LED emits red light when forward-biased [Fig. 56.9 (*a*)]. Fig. 56.9 (*b*) shows the schematic symbol of LED.



(*a*) (*b*)

Fig. 56.9

Uses

LEDs can be used in numeric displays such as seven segment display. A seven segment display unit is shown in Fig. 56.10. It is made by using a number of LEDs. By activating suitable combination of LEDs in this unit, any digit from 0 to 9 can be displayed by it. For example, switching ON the segments 1, 6, 7, 3, 4 displays the numeral 5.

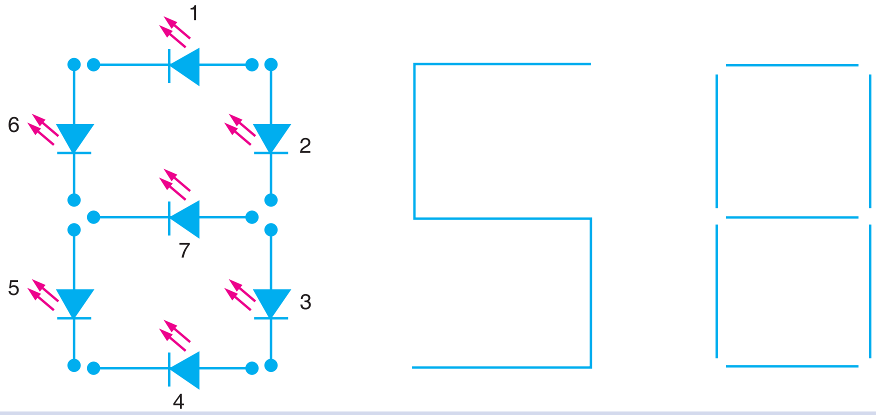


Fig. 56.10

When all the segments are ON, the digit displayed is 8. If only the segment 7 is OFF, the digit displayed is zero.

56.5 Liquid Crystal Display (LCD)

The action of the LCD of the field effect (twisted nematic) type is described below.

Liquid Crystal Display (L CD) Screen

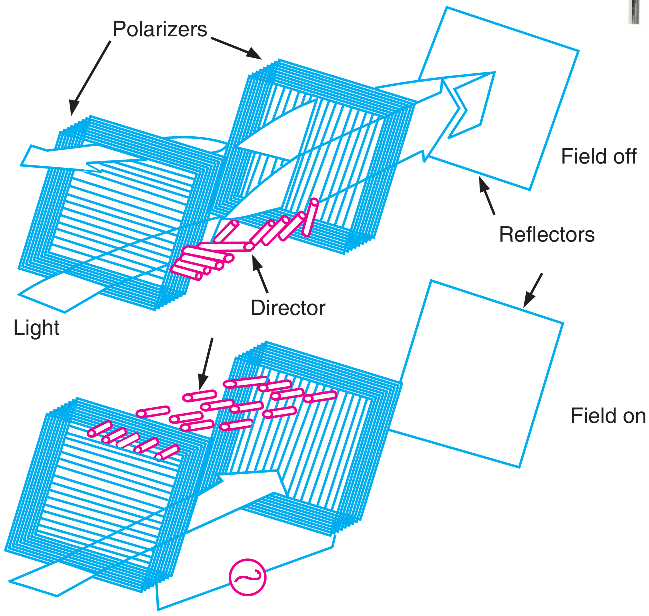


Fig. 56.11

In the twisted nematic configuration, the nematic phase is sandwiched between buffed glass plates that have transparent electrodes evaporated onto the buffed sides. The two plates are oriented with their buffing directions perpendicular to each other. As a consequence, the orientation of the molecules in nematic liquid crystal changes by 90° between the two plates (Fig. 56.11). The direction of the polarization of the light incident on the glass plate changes by 90°, because of the change in orientation of the molecules sandwiched between the two plates. Consequently when placed between crossed polaroids the light passes through. It gets reflected by a mirror and retraces its path. Thus the configuration looks bright when exposed to light. On applying an electric field between the plates, the molecules orient parallel to the field. Now the incoming light from the polariser cannot pass through the analyzer and hence the configuration looks dark when exposed to bright light. Hence the device can be used to display characters.

The display consists of two glass plates with a special liquid crystal (nematic fluid) in between [Fig. 56.12]. The bottom surface of the top plate has invisible semi-silvered shapes, where the segments or symbols are to be seen. The back glass plate is also metallized. A polariser (polaroid sheet) is kept over the top glass plate. Electrical contacts are made to the metallized back plate and each segment of the display in the top plate.

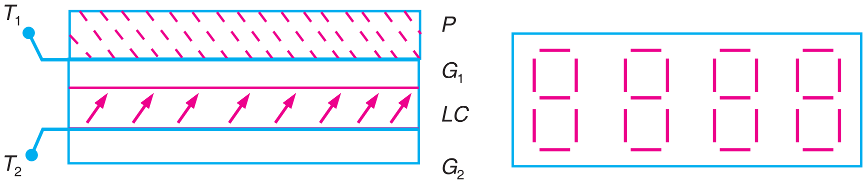


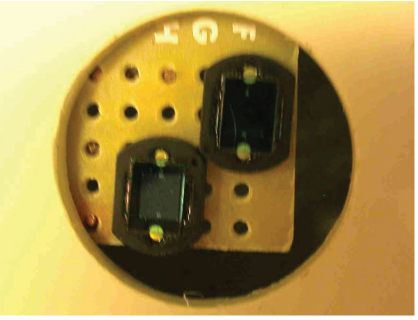
Fig. 56.12

When an electric field is applied across a segment and the back plate, the liquid crystal under that segment has its molecules oriented in a particular direction. The effect is that the liquid crystal and the polariser at the top are brought in 'crossed' position and there is extinction of light in that segment, when the field is present. Other segments, under no field, remain invisible.

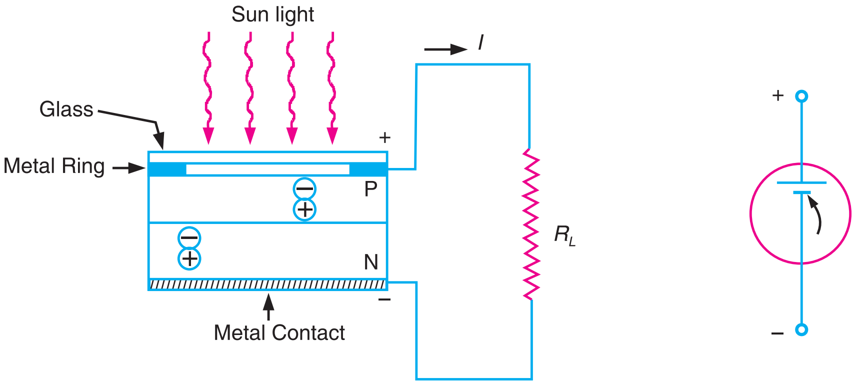
56.6 The Solar Cell

A *solar cell* is basically a *P-N* junction diode which converts solar energy (light energy) into electrical energy. In principle, a solar cell is nothing but a light emitting diode (LED) operating in reverse.

Common materials for solar cells include silicon, gallium Arsenide (Ga As), indium Arsenide (In As) and cadmium Arsenide (Cd As). The most common is silicon. For silicon, the band gap (the energy necessary to transfer an electron from the upper valence level to the conduction band) is 1.12 eV. The maximum theoretical efficiency of a solar cell depends on this band gap. For silicon, the maximum efficiency is 22%.



Solar Cell



(*a*) (*b*)

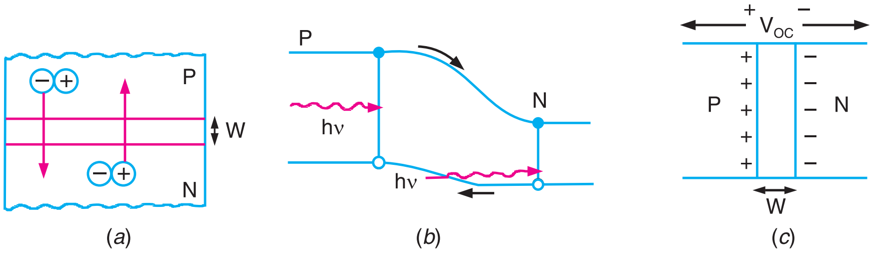
Fig. 56.13

Construction

A solar cell consists of *P-N* junction diode made of Si [Fig. 56.13 (*a*)]. Fig. 56.13 (*b*) gives the schematic symbol of a solar cell. The inward arrow indicates the incoming light. The *P-N* diode is packed in a can with glass window on top so that light may fall upon *P* and *N* type materials. The thickness of the *P*-region is kept very small so that electrons generated in this region can diffuse to the junction before recombination takes place. Thickness of *N*-region is also kept small to allow holes generated near the surface to diffuse to the junction before they recombine. A heavy doping of *P* and *N* regions is recommended to obtain a large photo voltage. A nickel plated ring is provided around the *P*-layer which acts as the positive output terminal. A metal contact at the bottom serves as the negative output terminal.

Working

The working of a solar cell may be understood from Fig. 56.14. When light radiation falls on a *P-N* junction diode, photons collide with valence electrons and impart them sufficient energy enabling them to leave their parent atoms. Thus electron-hole pairs are generated in both the *P* and *N* sides of the junction. These electrons and holes reach the depletion region *W* by diffusion [Fig. 56.14 (*a*)] and are then separated by the strong barrier field existing there. However, the minority carrier electrons in the *P*-side slide down the barrier potential to reach the *N*-side and the holes in the *N*-side move to the *P*-side [Fig. 56.14 (*b*)]. Their flow constitutes the minority current which is directly proportional to the illumination and also depends on the surface area being exposed to light.



(*a*) (*b*) (*c*)

Fig. 56.14

The accumulation of electrons and holes on the two sides of the junction [Fig. 56.14 (c)] gives rise to an *open circuit voltage* V_{oc} which is a function of illumination. The open-circuit voltage produced for a silicon solar cell is typically 0.6 volt and the short-circuit current is about 40 mA / cm^2 in bright noon day sun light.

Characteristics

Typical V-I characteristics of a solar cell, corresponding to different levels of illumination are shown in Fig. 56.15. Maximum power output is obtained when the cell is operated at the knee of the curve.

Relation between Load Current and Load Voltage

Holes and electrons produced in the junction region are swept to the *p*-type and *n*-type sides, respectively, as shown in Fig. 56.14. This produces a current I_s across the junction and also acts to charge the *p*-region positively and the *n*-type region negatively. Hence if there are no external connections to the junction, this forward bias causes a forward current

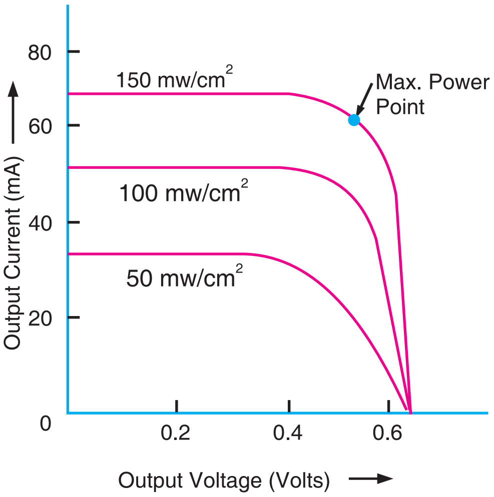


Fig. 56.15

I to flow. Under this condition, the forward current just balances the current I_g. When the p-and n-type sides are connected externally through an electrical load (Fig. 56.16), a portion of I_g flows in the external circuit.

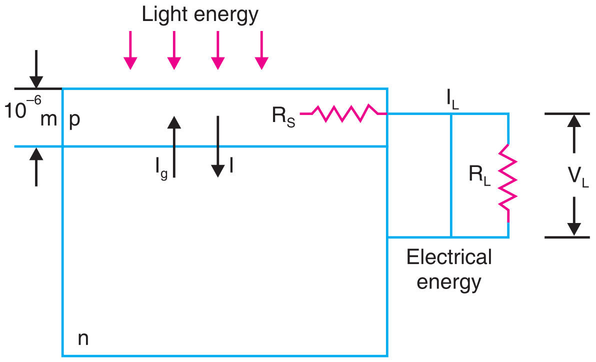


Fig. 56.16

The current I_g consists of that flowing across the junction and that flowing across the external load I_L.

$$I_g = I + I_L = I_o [\exp(eV/kT) - 1] + I_L \quad \text{...(1)}$$

The p-type region being made very thin, introduces an internal series resistance R_s in the solar battery which reduces its efficiency. The voltage drops in the circuit are

$$V = I_L R_S + V_L \quad \text{...(2)}$$

Eqs. (1) and (2) are combined to give the relation between load current and load voltage for a given I_g.

$$V_L = \frac{kT}{e} \log\left(1 + \frac{I_g - I_L}{I_O}\right) - R_S I_L \quad \text{...(3)}$$

Uses

1. Solar cells are used extensively in satellites and space vechicles to supply power to electronic and other equipments or to charge storage batteries.
2. Solar cells are used to power calculators and watches.
3. Solar cells are used to provide commercial electricity.

EXERCISE

1. Give the structure, symbol, volt-ampere characteristics and use of a photo diode.
2. Explain the phenomenon of photoconduction in semiconductors giving spectral response curve for Si and Ge. What is the effect of impurity on photoconductivity?
3. Explain with proper circuit diagram the construction, working and uses of a ‘Light Emitting Diode’. (*Rohilkhand,* 90)
4. Describe the action of a LCD. How does it compare with LED?
5. What is solar cell? Discuss the construction, working, I-V characteristics and uses of a solar cell.

CHAPTER 57

HALL EFFECT

AT A GLANCE

57.1 Hall Effect
57.2 Experimental Determination of Hall Coefficient

57.1 Hall Effect

When a magnetic field is applied perpendicular to a current carrying conductor, a voltage is developed across the specimen in a direction perpendicular to both the current and the magnetic field. This phenomenon is called the '*Hall effect*'. The voltage so developed is called the '*Hall voltage*'.

Hall voltage and Hall coefficient

Consider a slab of material subjected to an external electric field E_x along x-axis and a magnetic field H_z along z-axis (Fig. 57.1). As a result of the electric field, a current density J_x will flow in the direction of E_x . Let the current be carried by an electron of charge $-e$. The external magnetic field H_z applied perpendicular to current will exert a transverse magnetic deflecting force on the electron which causes the electron to drift downward to the lower edge of the specimen. Consequently, lower surface

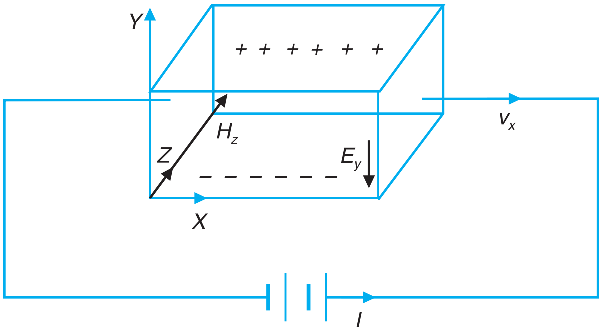


Fig. 57.1

of the slab collects a negative charge and upper surface a positive charge. This excess of negative charge on the lower edge and the corresponding excess of positive charge at the upper edge creates a transverse electric field E_y known as '*Hall electric field*' which opposes the transverse drifting of electrons. Ultimately an equilibrium is reached in which the force due to accumulation of electrons becomes equal to the magnetic force and so the flow of electrons stops. Thus the net force on electrons becomes zero, *i.e.*, $F_y = 0$.

The electric force on the electron having a charge $-e = -eE_y$

The force due to the magnetic field $= ev_x H_z$.

In the steady state $F_y = 0.$

$$F_y = -eE_y + e(v_x H_z) = 0$$

or $$E_y = v_x H_z \qquad \text{...(1)}$$

E_y is called the *Hall voltage* or *Hall field*.

Let n be the free electron density.

The current density is given by

$$J_x = -nev_x \qquad \text{...(2)}$$

$$v_x = -\frac{J_x}{ne}$$

$$\therefore \quad E_y = v_x H_z = -\frac{J_x H_z}{ne} = R_H J_x H_z \qquad \text{...(3)}$$

Here, $R_H = -\frac{1}{ne}$ is called the '*Hall coefficient*' for the substance of the bar.

The negative value of R_H indicates that current carriers are electrons. If the charge carriers are holes, R_H is positive.

Relation between R_H and μ

The mobility 'μ' is defined as the velocity acquired by the current carrying particle per unit electric field, *i.e.*,

$$\mu = \frac{v_x}{E_x} \qquad \therefore v_x = \mu E_x$$

Eq. (3) becomes $$E_y = \mu E_x H_z \qquad \text{...(4)}$$

Comparing Eqs. (3) and (4),

$$R_H J_x H_z = \mu E_x H_z$$

$$R_H = \mu \frac{E_x}{J_x}$$

$$\therefore \quad R_H = \mu \cdot \frac{1}{\sigma}$$

Here, σ is the electrical conductivity of the metal.

$$\therefore \quad \mu = R_H \sigma.$$

57.2 Experimental Determination of Hall Coefficient

A thin metallic strip, several mm wide and several cm long, is placed in x-direction and a magnetic field H_z is applied in z-direction (Fig. 57.2).

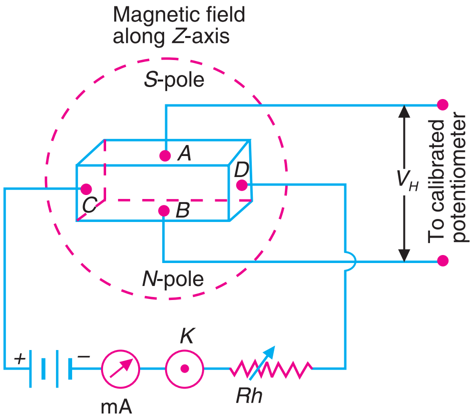
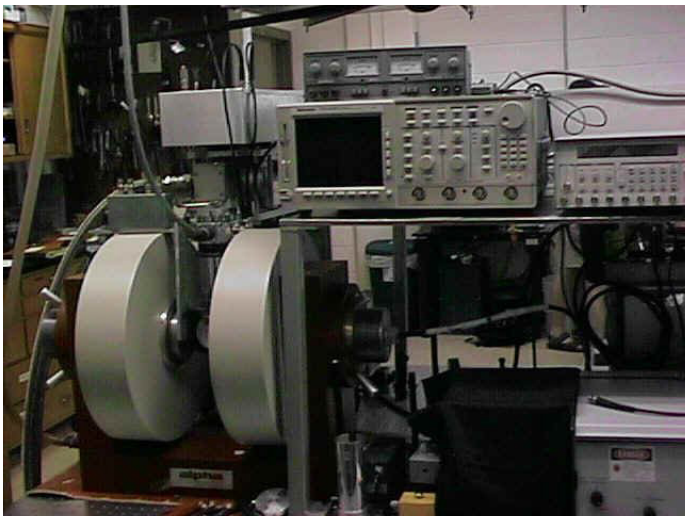


Fig. 57.2

A suitable current is passed in the specimen which can be adjusted by the rheostat. Two potential leads are placed between the points *A* and *B* which are connected to a sensitive calibrated potentiometer to measure the developed Hall voltage V_H. By measuring the Hall voltage, Hall-coefficient can be calculated as given below :



Experimental Determination of Hall Coefficient

The value of Hall electric field is given by

$$E_y = -\frac{J_x H_z}{ne}$$

Let *I* be the current flowing through the specimen.

Let *b* and *d* be respectively the breadth and thickness of the specimen.

The current density is given by

$$J_x = \frac{I}{b \times d}$$

$$E_y = -\frac{I}{nebd} \cdot H_z$$

or
$$E_y . d = V_H = -\frac{1}{neb} \cdot H_z = R_H \cdot \frac{I H_z}{b}$$

Hall coefficient,
$$R_H = \frac{V_H \cdot b}{I H_z}.$$

Thus, the value of Hall coefficient (R_H) can be calculated by knowing *I* and H_z and by measuring experimentally the Hall voltage V_H developed across the points *A* and *B* by means of a calibrated potentiometer.

Applications. The Hall effect measurements provide the following information about the solid:

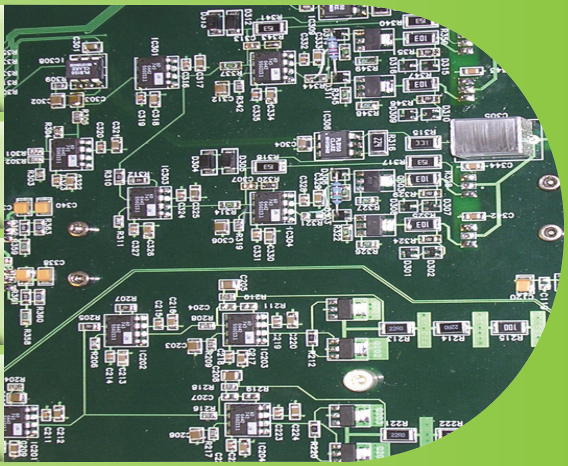
1. The sign (electrons or holes) of charge carrier is determined.
2. The carrier concentration (number of charge carriers per unit volume) is determined.
3. The mobility of charge carriers is measured directly.
4. We can decide whether a material is a metal, semiconductor or insulator.
5. From the knowledge of measured Hall voltage, the unknown magnetic field can be measured provided the value of Hall constant for the slab is known.

PART – VIII

ELECTRONICS

58. Electronics
59. Transistor Biasing
60. Electronics II
61. Hybrid Parameters
62. Transistor Amplifiers
63. Small Signal Amplifiers
64. FET Amplifier
65. Operational Amplifiers
66. Oscillators
67. Multivibrators and Logic Gates
68. Metal Oxide Semiconductor Field Effect Transistor (Mosfet)
69. IC Voltage Regulators
70. Digital Electronics
71. Half Adder and Full Adder
72. BJT Transistor Modeling
73. Unijunction Transistor (UJT)
74. Feedback Amplifiers

CHAPTER
58



ELECTRONICS

AT A GLANCE

58.1 Formation of p-n Junction Diode
58.2 Forward and Reverse Biasing of a Junction Diode
58.3 V-I Characteristics of a Junction Diode
58.4 The Zener Diode
58.5 Experiment to Study the Characteristics of the Zener Diode
58.6 Zener Diode as Voltage Regulator
58.7 P-N Junction Diode as Rectifier
58.8 Diode as a Half-Wave Rectifier
58.9 Mathematical Analysis
58.10 Full-Wave Rectifier
58.11 Mathematical Analysis
58.12 Bridge Rectifier
58.13 Filter Circuits
58.14 Junction Transistor
58.15 Common Emitter Characteristics of a Transistor
58.16 Common-Emitter Transistor Amplifier
58.17 Practical Circuit of Common Emitter Amplifier
58.18 A.C. Equivalent Circuit
58.19 Common-Base Transistor Amplifier
58.20 Oscillators
58.21 Condition for Self-Excitation (Barkhausen Criterion for Oscillations)
58.22 Hartley Oscillator
58.23 Colpitt's Oscillator
58.24 Tunnel Diode Operation
58.25 Tunnel Diode
58.26 Field-Effect Transistors
58.27 Drain Characteristics of an n Channel JFET
58.28 Principle of n-Channel FET Operation
58.29 Integrated Circuit (IC)
58.30 Fabrication of a Monolithic IC
58.31 Modulation and Detection
58.32 Amplitude Modulation
58.33 Power in an Amplitude Modulated Wave
58.34 Frequency Modulation
58.35 Demodulation or Detection
58.36 Block Diagram of AM Transmitting System
58.37 Block Diagram of AM Receivers—Superheterodyne Receiver
58.38 Characteristics of a Radio Receiver
58.39 Block Diagram of TV Transmitting System
58.40 Block Diagram of TV Receiving System

58.1 Formation of p-n Junction Diode

Junction diode is formed by placing a *p*-type crystal in contact with *n*-type crystal and subjecting to high pressure so that it becomes a single piece. The assembly so obtained is called *p*-*n* junction or junction diode or crystal diode. The surface of contact of *p* and *n*-type crystals is called junction. A *p*-*n* junction is shown in Fig. 58.1. The *p*-type region has (positive) holes as majority charge-carriers. Similarly, the *n*-type region has (negative) electrons as majority charge carriers.

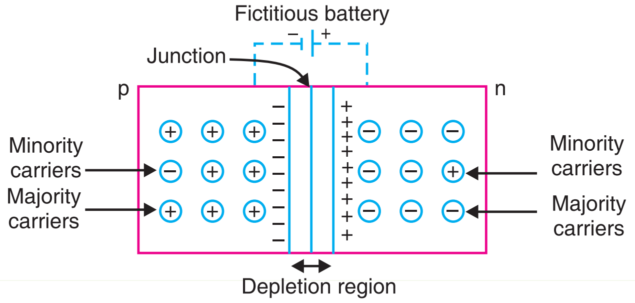
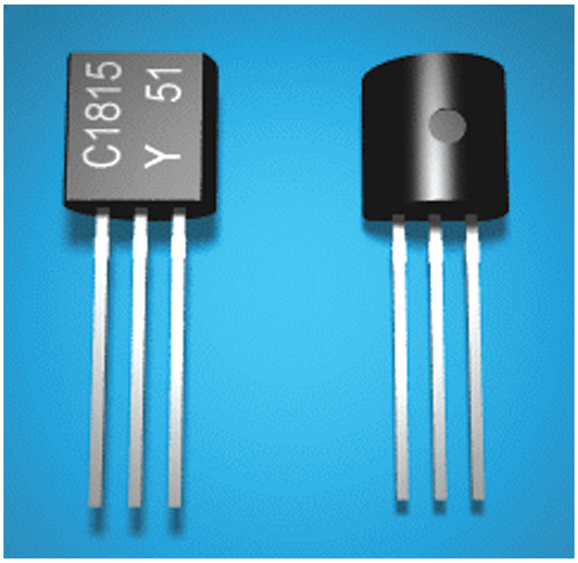


Fig. 58.1

In addition to these majority charge-carriers, there are a few minority charge-carriers in each region. The *p*-region contains a few electrons while the *n*-region contains a few holes. Thus we find that at the junction there is a decreasing hole concentration from left to right which makes the holes diffuse from *p*-side to *n*-side. Similarly electrons diffuse from right to left across the junction. Holes leaving and electrons entering the *p*-side make it negative. Similarly holes entering and electrons leaving the *n*-region make it positive. Thus, there is net negative charge on the *p*-side of the junction and net positive charge on the *n*-side. This produces an electric field across the junction as if a fictitious battery is connected with its positive terminal to *n* and negative terminal to *p* section. Equilibrium is established when the field becomes large enough to stop further diffusion of the majority charge-carriers. The field, however, helps the minority carriers to move across the junction.



p-n Junction Diode

The region on either side of the junction which becomes depleted (free) of the mobile charge-carriers is called the 'depletion region'. The thickness of this region is of the order of 10^{-6} m. The potential difference across the depletion region is called the 'potential barrier'.

58.2 Forward and Reverse Biasing of a Junction Diode

1. Forward Bias. When a battery is connected to the diode with *p*-section connected to positive pole and *n*-section to the negative pole, the junction diode is said to be forward biased (Fig. 58.2). If the forward bias is greater than the potential barrier, the majority carriers move towards the junction and cross it. The current which flows due to majority carriers is called forward current. It increases with forward bias.

2. Reverse Bias. When a battery is connected to junction diode with p-section connected to negative pole and n-section connected to the positive pole, the junction is said to be reverse biased (Fig. 58.3). When reverse bias is applied, the majority carriers do not cross the junction. However a very little amount of current flows due to the motion of minority carriers. This current is called reverse current. This current increases with increasing temperature.

Thus we see that the junction diode offers a low resistance for the current to flow in one direction (under forward bias) but a very high resistance in the opposite direction (under reverse bias). It thus acts as a rectifier.

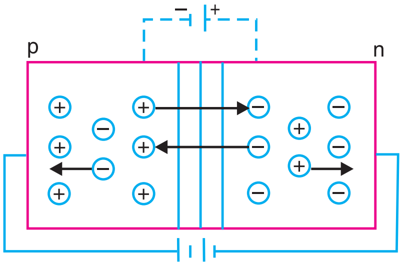


Fig. 58.2

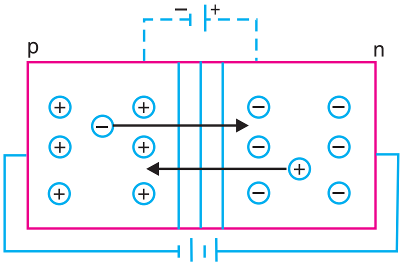


Fig. 58.3

58.3 V-I Characteristics of a Junction Diode

Graphs drawn between bias voltage and circuit current of a junction diode are called characteristics of the diode.

1. Forward Bias Characteristic. This is obtained by plotting a graph between forward bias voltage and circuit current. The circuit used is shown in Fig. 58.4. The milliammeter mA measures the current. The voltmeter V measures the p.d. across the diode. The forward voltage is gradually increased in steps and corresponding milliammeter readings are noted. A graph is then plotted between voltage and current (Fig. 58.5). Practically no current flows until the barrier voltage is overcome. Once the external voltage exceeds the barrier potential, the current increases rapidly, approximately exponentially.

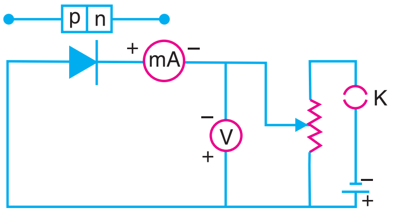


Fig. 58.4

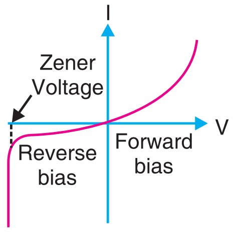


Fig. 58.5

2. Reverse Bias Characteristic. Connections are made as shown in Fig. 58.6. The reverse voltage is gradually increased in steps and corresponding microammeter readings are noted. A graph is then plotted between voltage and current (Fig. 58.5). With reverse bias the reverse current remains very small over a long range, increasing very slightly with increasing bias.

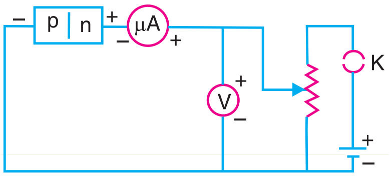


Fig. 58.6

***(a)* Avalanche breakdown.** If the reverse bias is made very high, the covalent bonds near the junction break down and a large number of electron-hole pairs are liberated. The reverse current then increases abruptly to a relatively large value. This is known as "*Avalanche breakdown*", and may damage the junction by excessive heat generated unless the current is limited by external circuit. This phenomenon is used in making **zener diodes.**

The maximum voltage that a junction diode can bear without breakdown is called zener voltage or reverse breakdown voltage.

(*b*) Dynamic resistance. Both the forward bias and reverse bias characteristics of the *p-n* junction do not obey Ohm's law. Therefore, the resistance offered by junction diode depends upon the applied voltage.

The dynamic resistance of junction diode is defined as the ratio of small change in voltage to the small change in current produced. It is also called a.c. resistance of the junction diode and is denoted by r_d.

Thus,
$$r_d = \frac{\Delta V}{\Delta I}.$$

The region of the characteristic curve, where dynamic resistance is almost independent of the applied voltage, is called the linear region of junction diode.

Example 1. *Can a p-n junction be formed for a given breakdown voltage ? Explain.*

Ans. Yes. It can be formed for a given breakdown voltage by suitably adding the doping agent.

58.4 The Zener Diode

The diode which operates in the reverse breakdown region with a sharp breakdown voltage is called a *Zener diode* (Fig. 58.7).

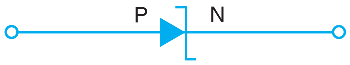
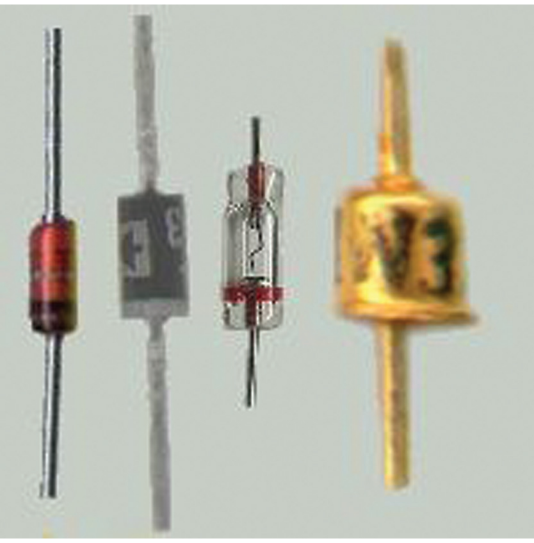


Fig. 58.7

It is an ordinary *P-N* junction diode except that it is properly doped to have a very sharp and almost vertical breakdown. It is exclusively operated under reverse bias conditions. It is designed to operate in breakdown region without damage. By adjusting the doping level it is possible to produce zener diodes with a breakdown voltage ranging from 2V to 800 V.

Zener diode primarily depends for its working on *Zener Effect.* In a heavily doped diode, the depletion region is very narrow. When the reverse bias voltage across the diode is increased, the electric field across the depletion region becomes very strong. When this field is ≈ 3×10^7 V/m, electrons are pulled out of the covalent bonds. A large number of electron-hole pairs are thereby produced. The reverse current rises steeply. This is *Zener effect.*

The external applied voltage accelerates the minority carriers in the depletion region. These carriers gain sufficient energy to ionise atoms by collision. The electrons produced thereby accelerate to sufficiently large velocities to be able to ionise other atoms. This creates a sort of chain reaction. The cumulative effect of this chain reaction is the *avalanche effect.*



Zener Diode

The zener effect is more prominent at breakdown voltages less than 4 V. The avalanche effect is more prominent above 6 V. Between 4 V and 6 V both effects are present. But, ordinarily, all diodes which are operated in the breakdown region of their reverse characteristics are called zener diodes.

58.5 Experiment to Study the Characteristics of the Zener Diode

***(i)* Forward Bias.** The circuit connections are made as shown in Fig. 58.8. The diode is given a forward biasing. The voltage is varied in regular steps of 0.1 volts by adjusting the rheostat.

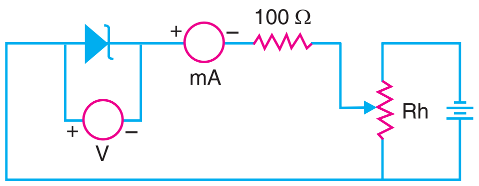


Fig. 58.8

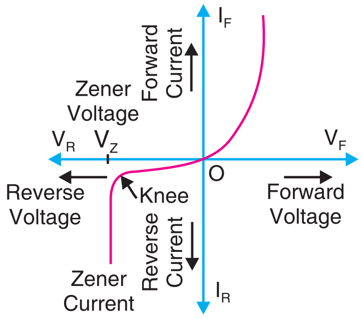


Fig. 58.9

The corresponding current is recorded using a milliammeter. The readings are tabulated. A graph is plotted taking voltage along the *x*-axis and current along the *y*-axis. Typical current-voltage characteristics for a zener diode are shown in Fig. 58.9. When forward biased, its characteristics are just that of an ordinary semiconductor diode.

***(ii)* Reverse Bias.** Connections are made as shown in Fig. 58.10. The voltage is varied in regular steps of 1 volt by adjusting the rheostat. The corresponding current is recorded using milliammeter and tabulated as shown. A graph of voltage against reverse current is drawn (Fig. 58.9).

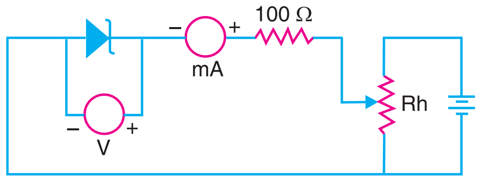


Fig. 58.10

Forward Bias		*Reverse Bias*	
V (volt)	I (mA)	V (volt)	I (mA)

When reverse biased, a small reverse saturation current flows through it. This current remains approximately constant until a certain critical voltage, called breakdown voltage, is reached. Beyond this voltage, the reverse current I_R increases sharply to a high value. This breakdown voltage V_z is called the *Zener voltage* and the reverse current as zener current.

V_z remains constant even when Zener current I_z increases considerably. This ability of a diode is called regulating ability which enables us to use zener diode for voltage regulation.

58.6 Zener Diode as Voltage Regulator

In a voltage regulated power supply unit, the output voltage is constant and it is independent of the variations of input supply voltage and load resistance. The circuit diagram of a Zener

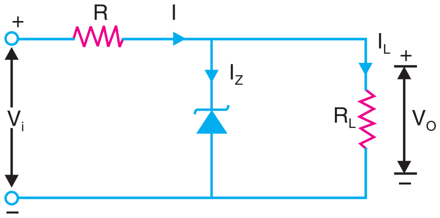


Fig. 58.11

diode voltage regulator is shown in Fig. 58.11. The unregulated dc is applied across the diode through a series resistor R, which limits the input current. The value of the series resitor R is so chosen that initially the diode operates in the breakdown region. The P-junction of the Zener diode is connected to the negative of the input voltage and N-junction to the positive. Thus the Zener diode is reverse biased. The output voltage V_0 remains essentially constant (equal to V_z) even though the input voltage V_i and the load resistance R_L may vary over a wide range.

Working. Let I be the current drawn from the supply source, I_z the current through the Zener diode and I_L that across the load resistance R_L.

Applying Kirchhoff 's laws, we get

$$I = I_z + I_L \qquad \text{...(1)}$$

$$V_0 = V_i - IR \qquad \text{...(2)}$$

and $$V_0 = I_L R_L \qquad \text{...(3)}$$

The variation in the output voltage may be due to two causes. First, the load current may vary. Second, the input voltage may vary.

Variation of load current : Suppose the load resistance R_L varies and the input voltage V_i remains constant. Since the output voltage V_0 tends to remain constant, Eq. (2) gives

$$\delta I = 0 \qquad (\because\ V_i \text{ and } R \text{ are constant})$$

Then Eq. (1) gives $$\delta I = \delta I_z + \delta I_L = 0$$

or $$\delta I_z = -\delta I_L.$$

Thus, *if the load resistance increases*, when the supply voltage is fixed, *the load current I_L decreases* and the Zener diode current I_z increases by an equal amount. Thus the *voltage V_0 across the load will tend to remain constant.*

Variation in input voltage : Now suppose that the load resistance R_L remains constant and supply voltage V_i varies. Since V_0 tends to remain constant, we get from Eq. (2),

$$\delta V_i = R\,\delta I$$

Also Eq. (3) gives, $$\delta I_L = 0 \qquad (\because\ R_L \text{ is constant})$$

∴ Eq. (1) gives, $$\delta I = \delta I_L$$

Thus when the supply voltage varies but the load resistance remains constant, the *total current I and the Zener current I_z change equally to keep the load current I_L constant.* Thus if total current I decreases by δI, the diode current I_z also decreases by the same amount, so that load current I_L remains constant and the *voltage V_0 across the load will tend to remain constant.*

Example 1. *Distinguish a Zener diode from an ordinary junction diode.*

Solution. 1. Ordinary junction diodes are operated within the breakdown voltage in reverse biased condition. Operation beyond breakdown voltage may damage them.

But a Zener diode is specially designed for operation beyond breakdown voltage. This breakdown voltage is called the *Zener breakdown potential.*

2. Ordinary diode is usually used for rectification while Zener diode is used for voltage regulation.

Example 2. *A 9 volt stabilized voltage supply is required to run a car stereo system from the car's 12 volt battery. A Zener diode with $V_z = 9$ V and $P_{max} = 0.27$ W is used as voltage regulator. Find the value of the series resistor. If the load resistance is 450 Ω, find the diode current.*

Solution. Let I – current drawn from the battery
I_z – current through Zener diode
I_L – current through load resistance (Fig. 58.12).

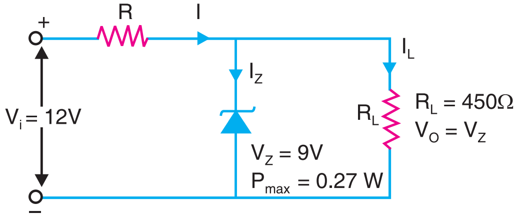


Fig. 58.12

Voltage drop across $R = V_i - V_L = V_i - V_Z = 12 - 9 = 3$ V.
Potential difference across $R_L = V_Z = V_0 = 9$ V.

$$\therefore \quad \text{Load current } I_L = \frac{\text{p.d. across } R_L}{\text{load resistance}} = \frac{9}{450} = 0.02 \text{ A}.$$

$$\text{Zener current } I_Z = \frac{P_{\max}}{V_{\max}} = \frac{0.27}{9} = 0.03 \text{ A}\cdot$$

∴ The current drawn from the battery

$$I = I_Z + I_L = 0.03 \text{ A} + 0.02 \text{ A} = 0.05 \text{ A}.$$

∴ The value of series resistor,

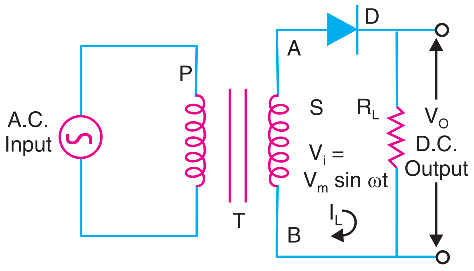
$$R = \frac{\text{voltage drop across } R}{\text{current through } R} = \frac{3}{0.05} = 60\,\Omega\cdot$$

58.7 P-N Junction Diode as Rectifier

Rectification is the process in which ac is converted into dc. The device which is used for rectification is called a rectifier. A junction diode allows a current to flow through it when it is forward biased. This property of diode is used for rectification. A half-wave rectifier is one which converts a.c. voltage into a pulsating voltage using only one half cycle of the applied a.c. voltage. A full wave rectifier is one which converts a.c. voltage into a pulsating voltage using full cycle of the applied a.c. voltage.

58.8 Diode as a Half-Wave Rectifier

Construction. Fig. 58.13 shows the circuit for a half-wave rectifier. T is a transformer. The primary of the transformer is connected to the ac mains. The diode D is connected across the secondary in series with a load resistance R_L.



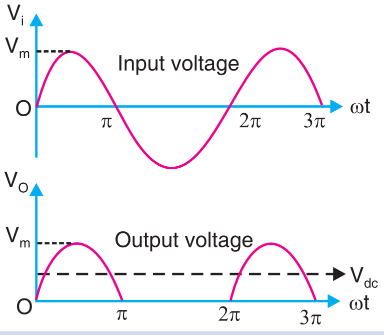


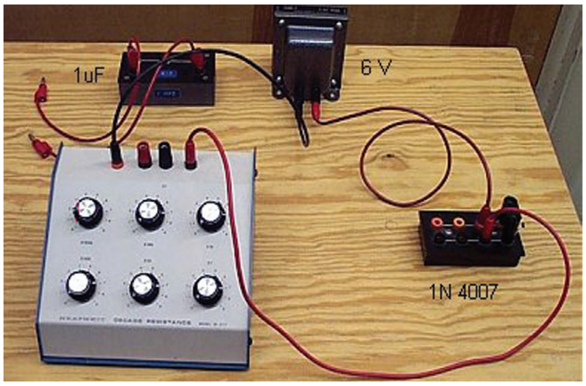
Fig. 58.13

Fig. 58.14

Working. The primary of the transformer is connected to the ac mains. An ac voltage will be induced across the secondary. This voltage can be represented by

$$V_i = V_m \sin \omega t.$$

Fig. 58.14 shows the variation of this input voltage with time. V_m is the peak value.



Half-wave Rectifier Circuit

During the first half cycle of a.c., one end of the secondary, say A, becomes positive. Then the diode is forward biased. Hence current flows through the load R_L in the direction of arrows (Fig. 58.13). The diode offers very little resistance when forward biased. Hence the p.d. across it is very small. The voltage across the load R_L is therefore practically the same as that across the secondary of the transformer, *i.e.*, V_i. During the next half cycle, the end A becomes negative. The diode is now reverse biased. Therefore, no current flows through the load R_L. The voltage across the load is zero. The shape of the output voltage is shown in Fig. 58.14. This voltage in not a perfect dc. But it is unidirectional.

58.9 Mathematical Analysis

The input voltage applied to the diode is given by

$$V_i = V_m \sin \omega t \qquad \text{...(1)}$$

Then, the instantaneous output current through the load resistance R_L is given by

$$I_L = I_m \sin \omega t \qquad \text{when } 0 \le \omega t \le \pi \qquad \text{...(2)}$$

and

$$I_L = 0 \qquad \text{when } \pi \le \omega t \le 2\pi \qquad \text{...(3)}$$

$$I_m = \frac{V_m}{R_L + R_f}$$

Here,

I_m = peak value of the current

R_f = dynamic forward resistance of the diode.

(*i*) D.C. (average) value of output current

The average d.c. current over one complete cycle is given by

$$I_{av} = I_{dc} = \frac{1}{2\pi}\int_0^{2\pi} I_L \, d(\omega t)$$

$$= \frac{1}{2\pi}\left[\int_0^{\pi} I_m \sin(\omega t)\, d(\omega t) + \int_0^{2\pi} 0 \times d(\omega t)\right]$$

$$= \frac{I_m}{2\pi}[-\cos \omega t]_0^{\pi}$$

$$\therefore \quad I_{dc} = \frac{I_m}{\pi} = \frac{1}{\pi}\left[\frac{V_m}{R_f + R_L}\right] \qquad \text{...(4)}$$

The dc voltage developed across the load R_L is given by

$$V_{dc} = I_{dc} \times R_L = \frac{I_m}{\pi} R_L \qquad \text{...(5)}$$

(*ii*) D.C. power output. The dc power output across the load R_L is

$$\text{D.C. power output} = P_{dc} = I_{dc}^2 \cdot R_L = \frac{I_m^2}{\pi^2} \cdot R_L \qquad ...(6)$$

(*iii*) R.M.S. (effective) value of output current

The root mean square value of the current, by definition, is given by

$$I_{rms} = \left[\frac{1}{2\pi}\int_0^{2\pi} i_L^2 \, d(\omega t)\right]^{1/2}$$

$$= \left[\frac{1}{2\pi}\left\{\int_0^{\pi} i_m^2 \sin^2 \omega t \, d(\omega t) + \int_{\pi}^{2\pi} 0 . d(\omega t)\right\}\right]^{1/2}$$

$$= \left[\frac{1}{2\pi}\int_0^{\pi} i_m^2 \sin^2 \omega t \, d(\omega t)\right]^{1/2}$$

$$\therefore \quad I_{rms} = \frac{I_m}{2} \qquad ...(7)$$

(*iv*) A.C. power input : The power supplied to the circuit from the ac source is given by

$$P_{AC} = I_{rms}^2 (R_f + R_L) = \frac{I_m^2}{4}(R_f + R_L) \qquad ...(8)$$

(*v*) Rectifier efficiency : It is defined as the ratio of dc output power to the total ac power supplied to the rectifier.

$$\eta = \frac{\text{D.C. power output}}{\text{A.C. power input}}$$

$$= \frac{I_m^2 \, R_L / \pi^2}{I_m^2 \, (R_f + R_L)/4} = \frac{4}{\pi^2}\left[\frac{R_L}{R_L + R_f}\right]$$

$$\eta = \frac{0.406}{1 + R_f / R_L}$$

Efficiency is maximum when $R_f << R_L$.

Theoretical maximum efficiency $\eta = 0.406 = 40.6\%$.

(*vi*) Ripple factor : The ripple factor is the ratio of r.m.s. value of A.C. component to the D.C. component in the rectifier output, *i.e.*,

$$\gamma = \frac{\text{r.m.s. value of A.C. component of output voltage}}{\text{D.C. component of output voltage}}$$

$$\gamma = \frac{V_{ac}}{V_{dc}} = \frac{I_{ac}}{I_{dc}}$$

The effective (r.m.s.) value of total load current is given by,

$$I_{rms} = \sqrt{I_{dc}{}^2 + I_{ac}{}^2}$$

$$I_{ac} = \sqrt{I_{rms}{}^2 - I_{dc}{}^2}$$

$$\frac{I_{ac}}{I_{dc}} = \frac{1}{I_{dc}}\sqrt{I_{rms}{}^2 - I_{dc}{}^2}$$

$$\gamma = \sqrt{\left(\frac{I_{rms}}{I_{dc}}\right)^2 - 1}$$

For a half-wave rectifier,

$$\frac{I_{rms}}{I_{dc}} = \frac{I_m/2}{I_m/\pi} = \frac{\pi}{2}\cdot$$

$$\gamma = \sqrt{\frac{\pi^2}{4} - 1} = 1.21\,.$$

This indicates that the amount of A.C. component present in the output of a half-wave rectifier is 121% of D.C. output voltage. The half-wave rectifier is therefore a poor converter of ac into dc.

(*vii*) Peak inverse voltage. Peak inverse voltage (PIV) is defined as the *maximum reverse voltage which the rectifier has to withstand during the non-conducting period.* Thus for a half-wave rectifier,

$$PIV = V_m.$$

(*viii*) Voltage regulation. Voltage regulation is the ability of a rectifier to *maintain a specific output voltage irrespective of the variation in the load resistance.* In a half-wave rectifier,

$$I_{dc} = \frac{I_m}{\pi} = \frac{V_m}{\pi(R_L + R_f)}$$

or $$I_{dc} R_L = \frac{V_m}{\pi} - I_{dc} R_f$$

$$\therefore \quad V_{dc} = \frac{V_m}{\pi} - I_{dc} R_f\,.$$

When $I_{dc} = 0$, V_{dc} has its maximum value (V_m / π). As I_{dc} increases, V_{dc} decreases linearly. Therefore, *voltage regulation of a half-wave rectifier is poor.*

Example 1. *The applied input a.c. power to a half-wave rectifier is* 100 *watt. The d.c. output power obtained is* 40 *watt. Calculate the rectification and power efficiency of the above circuit. What happens to the remaining* 60 *watt*?

Solution. The rectification efficiency is given by

$$\eta = \frac{\text{d.c. output}}{\text{a.c. input}} \times 100 = \frac{40}{100} \times 100 = 40\%$$

Half-wave rectifier rectifies only half the cycle of sinusoidal wave. The input power is only 50 W.

$$\text{Power efficiency} = \frac{\text{output power}}{\text{input power}} \times 100 = \frac{40}{50} \times 100 = 80\%\cdot$$

Example 2. *The turns ratio of a transformer used in a half-wave rectifier is* 12 : 1. *The primary of the transformer is connected to* 220 *V,* 60 *Hz mains. If the diode resistance on forward bias is zero, calculate the d.c. voltage across the load. Also find the Peak Inverse Voltage.*

Solution. Turns ratio of a transformer is given by

$$\frac{N_p}{N_s} = \frac{V_p}{V_s}$$

$$\therefore \quad V_s = \frac{N_s}{N_p}\cdot V_p = \frac{1}{12} \times 220 = 18.33\,\text{V}$$

∴ Maximum voltage across secondary

$$= \sqrt{2}\,V_s = \sqrt{2} \times 18.33 = 25.92\,\text{V}$$

$$\therefore \qquad \text{D.C. output voltage} = \frac{V_m}{\pi} = \frac{25.92}{3.14} = 8.252\,\text{V}$$

$$\text{Peak inverse voltage} = 25.92\ \text{V.}$$

58.10 Full-Wave Rectifier

In a full-wave rectifier both halves of the input-cycle are used. There are two types of full-wave rectifiers: (1) Centre tapped full-wave rectifier, and (2) Bridge rectifier.

Centre tapped full-wave rectifier. A full wave rectifier circuit consists of two diodes D_1 and D_2 connected to the secondary of the step-down transformer. The input A.C. signal is fed to the primary of the transformer (Fig. 58.15).

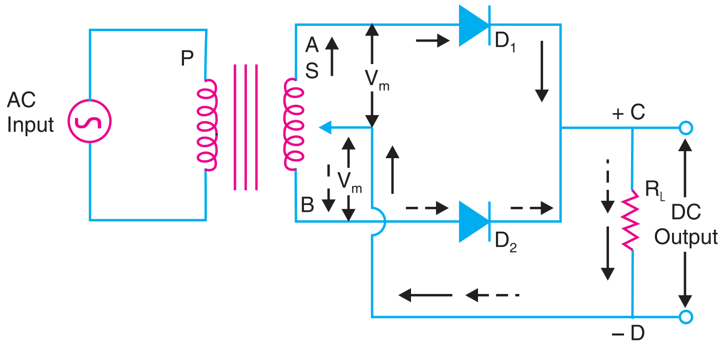


Fig. 58.15

Working. During the positive half-cycle of the secondary voltage, one end of the secondary, say A, becomes positive and end B becomes negative. So the diode D_1 is forward biased, and diode D_2 is reverse biased. As a result of this, the diode D_1 conducts current whereas the diode D_2 does not conduct. Current through the load resistance flows from C to D producing output voltage V_0. The current is shown by solid arrows.

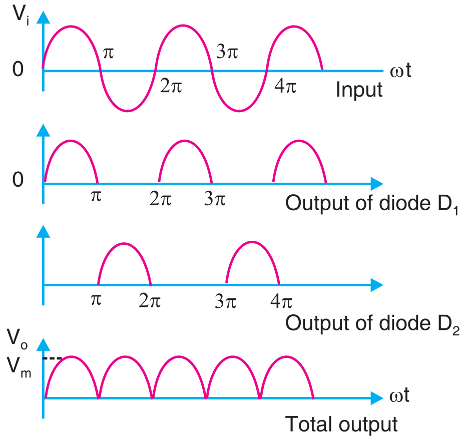


Fig. 58.16

Full-wave Rectifier

During the negative half cycle of *AC* input, end *A* becomes negative and end *B* positive. So the diode D_1 is reverse biased and the diode D_2 is forward biased. As a result, the diode D_1 does not conduct and D_2 conducts current. Again current flows from *C* to *D* through the load resistance R_L producing output voltage V_0. The current is shown by the dotted arrows.

Thus, during both the half cycles, current flows through the load in the same direction. The output voltage is developed across the load R_L during the entire cycle. It is a pulsating D.C. voltage containing both A.C. and D.C. components. The input and the rectified output wave-forms are shown in Fig. 58.16.

58.11 Mathematical Analysis

Let the diodes D_1 and D_2 be identical and have the same dynamic resistance R_f. At any instant, let the magnitudes of AC voltages applied to the diodes be each equal to $V_i = V_m \sin \omega t$. V_m is the peak input voltage.

Let R_f = dynamic forward resistance of the diode.

The current pulses in the two diodes are given by

$$i = \begin{Bmatrix} I_m \sin \omega t & \text{for } 0 < \omega t < \pi \\ -I_m \sin \omega t & \text{for } \pi < \omega t < 2\pi \end{Bmatrix} \qquad \text{...(1)}$$

Here, $$I_m = \frac{V_m}{R_f + R_L}.$$

(*i*) D.C. (average) value of output current. The output dc current I_{dc} is given by

$$I_{dc} = \frac{1}{2\pi}\int_0^{2\pi} i\, d(\omega t)$$

$$= \frac{1}{2\pi}\left[\int_0^{\pi} I_m \sin \omega t\, d(\omega t) + \int_{\pi}^{2\pi} -I_m \sin \omega t\, d(\omega t)\right]$$

$$= \frac{I_m}{2\pi}\left[[-\cos \omega t]_0^{\pi} + [\cos \omega t]_{\pi}^{2\pi}\right] = \frac{I_m}{2\pi}[2+2]$$

$$\therefore \quad I_{dc} = \frac{2I_m}{\pi}. \qquad \text{...(2)}$$

(*ii*) R.M.S. (effective) value of load current. The r.m.s. value of total output current is given by

$$I_{rms} = \sqrt{\frac{1}{2\pi}\int_0^{2\pi} i^2 d(\omega t)}$$

$$= \left[\frac{1}{2\pi}\left\{\int_0^{\pi} I_m^2 \sin^2 \omega t.d(\omega t) + \int_{\pi}^{2\pi} I_m^2 \sin^2 \omega t.d(\omega t)\right\}\right]^{1/2}$$

$$\therefore \quad I_{rms} = \frac{I_m}{\sqrt{2}}. \qquad \text{...(3)}$$

(*iii*) Power supplied to the circuit. The a.c. power input to the rectifier from the supply is given by

$$P\propto = I_{rms}^2 (R_f + R_L) = \frac{(R_f + R_L) I_m^2}{2}. \qquad \text{...(4)}$$

(*iv*) Average power supplied to the load R_L. The d.c. power output across the load R_L is given by

$$P_{dc} = I_{dc}^2 R_L = \frac{4I_m^2 R_L}{\pi^2}. \qquad \text{...(5)}$$

(*v*) Rectifier efficiency. In a rectifier, the useful power output is the d.c. power which is developed across the load R_L. Therefore, efficiency

$$\eta = \frac{\text{d.c. power supplied to the load}}{\text{Total input A.C. power}} \times 100\%$$

$$= \frac{P_{dc}}{P_{ac}} \times 100\% = \frac{4I_m^2 R_L / \pi^2}{(R_f + R_L) I_m^2 / 2} \times 100\%$$

[Using Eqs. (4) and (5)]

$$\therefore \qquad \eta = \frac{81.2}{1 + \frac{R_f}{R_L}} \% \qquad \text{...(6)}$$

Thus, the rectification efficiency of a full-wave rectifier is double that of a half-wave rectifier under identical conditions.

The maximum possible efficiency of a full-wave rectifier is 81.2 % when $R_f << R_L$.

(*vi*) Ripple factor. The ripple factor γ is given by

$$\gamma = \sqrt{\left(\frac{I_{rms}}{I_{dc}}\right)^2 - 1} = \sqrt{\left(\frac{I_m / \sqrt{2}}{2I_m / \pi}\right)^2 - 1}$$

$$\therefore \qquad \gamma = \mathbf{0.482}$$

The ripple factor of a full-wave rectifier is 0.482 and is much smaller than that of half-wave rectifier. Hence, in actual practice, a full-wave rectifier is preferred to a half-wave rectifier.

(*vii*) Peak inverse voltage. *PIV is the maximum reverse voltage which the rectifier has to withstand during the non-conduction period.* Suppose the diode D_1 is conducting and D_2 non-conducting.

The reverse voltage across diode D_2

$$\text{PIV} = \text{PD across } R_L - (-V_m) = V_m + V_m = 2V_m$$

Thus in a full-wave rectifier, PIV across each diode is two times the maximum transformer voltage measured from the centre-tap to either end.

Example 1. *A full-wave rectifier supplies d.c. to a load of* $1 k\Omega$. *If the a. c. voltage applied to the diodes is* 200-0-200 *volts* (*rms*), *calculate:* (*a*) *Average d.c. voltage*, (*b*) *Average d.c. current*, (*c*) *Ripple voltage* (*rms*), *neglecting the diode resistance.*

Solution. The peak value of dc current is given by

$$I_m = \frac{V_m}{R_L} = \frac{V_{rms}\sqrt{2}}{R_L} = \frac{200\sqrt{2}}{10^3} = 0.2828 \text{ A}$$

The average d.c. current is given by

$$I_{dc} = \frac{2I_m}{\pi} = \frac{2 \times 0.2828}{\pi} = 0.18 \text{ A}$$

Average d.c. voltage is $V_{dc} = I_{dc} \times R_L = 0.18 \times 10^3 = 180$ V

Ripple factor, $\gamma = \sqrt{\frac{V^2{}_{rms}}{V^2{}_{dc}} - 1} = \sqrt{\left(\frac{200}{180}\right)^2} - 1 = \mathbf{0.4843}$

Ripple voltage $\quad V_{rms} = \gamma \times V_{dc} = 0.4843 \times 180 = 87.17 \text{ V}$.

Example 2. *In a full-wave rectifier the load resistance is 1 kΩ. The forward dynamic resistance of each diode is 10 Ω. The voltage across half of the secondary winding is 220 sin 200 t. Find the (a) Peak value of the current. (b) Average dc value of the current. (c) The rms value of the current. (d) The rectification efficiency.*

Solution. The instantaneous value of the A.C. input voltage is given by

$$V = V_m \sin \omega t$$

where V_m is peak value of the A.C. input voltage.

$$\text{Peak value of current, } I_m = \frac{V_m}{(R_f + R_L)} = \frac{220}{(10 + 1000)} = \mathbf{0.2178\ A}$$

The average d.c. current of full-wave rectifier is given by

$$I_{dc} = \frac{2I_m}{\pi} = \frac{2 \times 0.2178}{\pi} = \mathbf{0.1387\ A}$$

$$\text{r.m.s. value of current, } \quad I_{rms} = \frac{I_m}{\sqrt{2}} = \frac{0.2178}{\sqrt{2}} = \mathbf{0.154\ A}$$

$$\text{Rectifier efficiency, } \eta = \frac{\text{d.c. power output}}{\text{a.c. power input}}$$

$$\eta = \frac{I^2_{dc}\ R_L}{I^2_{rms}\,(R_f + R_L)} = \frac{(0.1387)^2 \times 1000}{(0.154)^2 \times (10 + 1000)} = \mathbf{0.803}$$

Rectifier efficiency is 80.3 %.

Example 3. *Is the frequency content of the output of a half-wave rectifier and a full-wave rectifier the same? Explain.*

Solution. No. Full-wave rectifier rectifies both the half-cycles. Hence the frequency of the output current is double that of the input current. But in a half-wave rectifier only one-half of the input signal is rectified. Hence the frequency of the output current is the same as that of the input current.

58.12 Bridge Rectifier

The circuit is shown in Fig. 58.17. The diodes D_1, D_2, D_3 and D_4 are arranged in the form of Wheatstone Bridge network. The two opposite ends A and C of the network are connected to the ends S_1 and S_2 of the secondary of transformer T. The ends B and D are connected to the load resistance R_L. The primary P of the transformer is connected to the ac mains.

When an AC voltage is applied to the primary, at some instant the positive half of the input cycle passes through the secondary, keeping the point A positive and C negative. Diodes D_1 and D_3 conduct and a current flows in the direction $ABR_L\ DC\ S_2\ S_1\ A$.

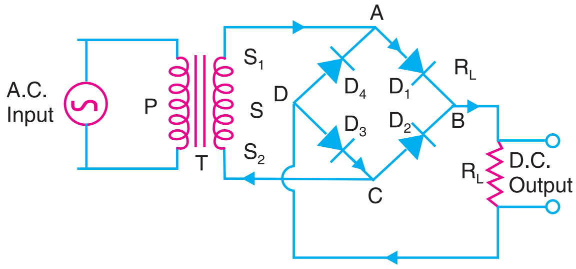
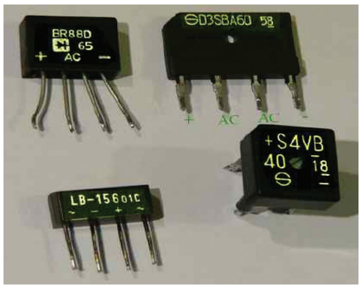


Fig. 58.17



Bridge Rectifier

At the same time D_2 and D_4 will not be conducting, since they are reverse biased. During the next half cycle, the point *A* is negative and *C* is positive. Therefore in this case diodes D_2 and D_4 conduct and current flows in the direction $CBR_L\ DAS_1\ S_2\ C$. But now D_1 and D_3 will not be conducting. Therefore during both the halves of the input cycle, current flows through the load R_L in the same direction. Thus a *DC* output is developed across R_L and we have full wave rectification.

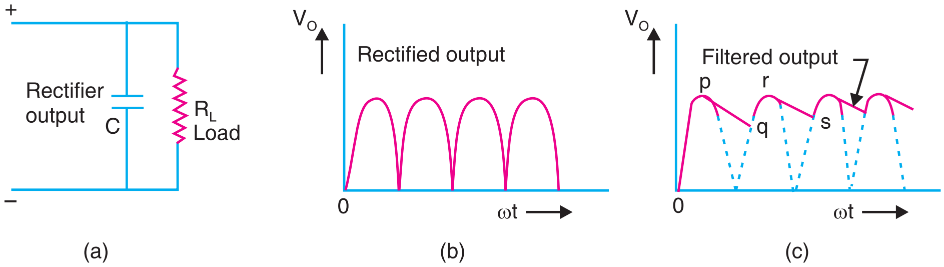
58.13 Filter Circuits

The output of a rectifier is a unidirectional voltage (or current), which is pulsating. It can be considered as consisting of a pure d.c. voltage with superposed a.c. components, known as "ripples". In electronic circuits we require a constant d.c. supply. Therefore, the ripples must be removed or minimised from the rectifier output. This can be done by introducing an electrical circuit consisting of choke coils and capacitors between the rectifier and the load. Such a circuit is known as a 'filter circuit'. There are several types of filters. Here we study a few of them.

***(i)* Shunt capacitor filter.** The simplest type of filter is a shunt capacitor connected across the rectifier output and the load (Fig. 58.18 *a*). It is an effective way of filtering the a.c. components from the output of the rectifier. The capacitor *C* is so chosen that its reactance ($1/\omega C$) at the frequency of a.c. mains is very small as compared to the load R_L. Then the a.c. components find a low reactance shunt path through the capacitor and are mostly bypassed. Thus ripples are filtered from the output voltage.

The pulsating output voltage of the rectifier, in the absence of capacitor C, is shown in Fig. 58.18 (*b*). The filtering action of the capacitor is indicated in Fig. 58.18 (*c*). The capacitor is charged almost equal to the peak voltage of the rectifier output at the end of the first quarter-cycle (point *p*).

During the subsequent half-cycle, when the rectifier output drops to zero and then rises again, the capacitor discharges slowly (exponentially) through the load. This discharge is represented by *pq*. At *q* the capacitor begins to be recharged to the peak voltage of the rectifier output (point *r*), and then discharges again to the point *s*. This cycle is repeated. Thus the voltage across the capacitor (and hence across the load) varies as shown by the solid curve in Fig. 58.18 (*c*). The variations are now much less than in the rectifier output (Fig. 58.18 *b*). If the time-constant (*CR*) of the filter is large, the capacitor discharges quite slowly and the average dc voltage across the load remains nearly constant.



(a) (b) (c)

Fig. 58.18

For a full-wave rectifier with capacitor filter, the ripple factor is given by,

$$\gamma = \frac{1}{4\sqrt{3}\,f\,C\,R_L}$$

where *f* is the supply frequency.

***(ii)* Series inductor filter.** In this filter, an inductor *L* (or choke) is placed in series with the load resistor R_L (Fig. 58.19).

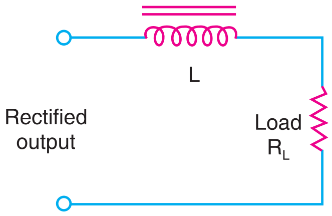


Fig. 58.19

An inductor opposes any change in the current flowing through it. The output of the rectifier (ac + dc) is fed to the choke. The choke offers a high reactance to the ac component and a small resistance to the dc component. Consequently, the amplitude of the ac component in the output voltage is considerably reduced.

For a full-wave rectifier with series inductor filter, the ripple factor is given by

$$\gamma = 0.236\left(\frac{R_L + R_C}{\omega L}\right)$$

where R_C is the resistance of the inductor.

***(iii)* The L-section or Inductor Input Filter.** In this type of filter circuit, choke L is connected in series with the output of a rectifier and capacitor C is connected parallel to the load R_L (Fig. 58.20). The series inductor L readily passes the d.c. components from the rectifier output but offers a high impedance (ωL) to the ac pulsations. Any pulsations in current that remains after it passes through L are largly by passed by the capacitor which offers a low impedance ($1/\omega C$) to them, but an infinite resistance to dc. The current flowing in the load is therefore very nearly steady.

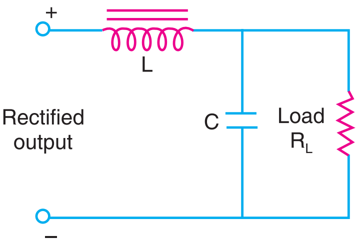


Fig. 58.20

In a full-wave rectifier with inductor input filter, ripple factor

$$\gamma = \frac{0.47}{4\omega^2\ LC - 1}.$$

Here, ripple factor is independent of load resistance. It is used in large current power supply.

***(iv)* The π-section (capacitor-input) filter.** This type of filter is shown in Fig. 58.21. It may be considered to be made up of the following two components :

(*i*) Capacitor filter formed by capacitor C_1.

(*ii*) Inductor input filter formed by choke L and capacitor C_2.

The capacitor C_1 is periodically charged to almost the peak value of the rectifier output. Between successive charging pulses, the voltage across C_1 falls off somewhat due to its slow discharge through the L-section filter and the load, but remains very near the peak value. The remaining pulsations in the current are opposed by the series inductor L and bypassed to ground by capacitor C_2. In this way the ripple in the output voltage is reduced to a large extent.

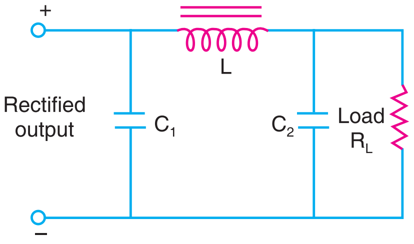


Fig. 58.21

In a full-wave rectifier with π filter, ripple factor

$$\gamma = \frac{1}{32\sqrt{2}\,\pi^3 f^3 C_1 C_2 L R_L}.$$

BIPOLAR JUNCTION TRANSISTOR (BJT)

58.14 Junction Transistor

The junction transistors are of two types : *p-n-p* transistor and *n-p-n* transistor. Fig. 58.22 (*a*) shows an *n-p-n* transistor and its symbol. Fig. 58.22 (*b*) shows a *p-n-p* transistor and its symbol. We consider only the *n-p-n* transistor. A transistor has the following sections :

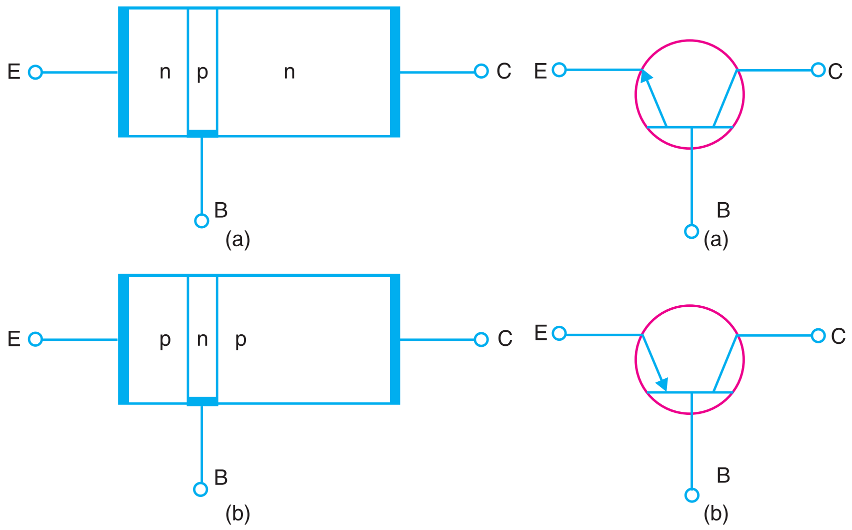


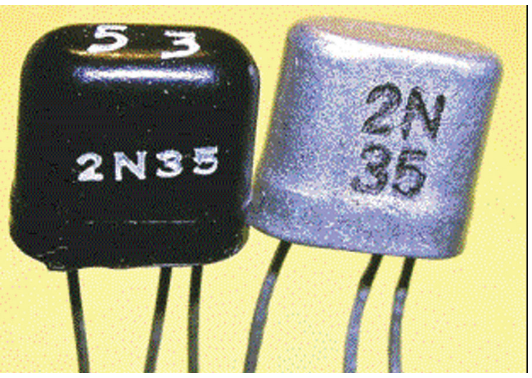
Fig. 58.22

(*i*) Emitter. This forms the left hand section or region of the transistor. The main function of this region is to supply majority charge carriers (either electrons or holes) to the base. Hence it is more heavily doped in comparison to other regions. Electrons are the charge-carriers within the *n-p-n* transistor, whereas holes are the charge-carriers within a **p-n-p transistor.**

(*ii*) Base. The middle section of the transistor is called base. This is very lightly doped and is very thin (10^{-6} m) as compared to either emitter or collector so that it may pass most of the injected charge carriers to the collector.

(*iii*) Collector. The right hand side of the transistor is called collector. The main function of the collector is to collect majority charge carriers through the base. This is moderately doped.

The junction between the emitter and base is called emitter junction. The junction between the collector and the base is called collector junction. In normal operation of a transistor, the emitter-base junction is forward biased while the collector-base juction is reverse biased.



Junction Transistor.

Working of an *n-p-n* Transistor. The emitter-base junction is forward biased by the battery V_{EB} (Fig. 58.23). The collector-base junction is reverse biased by the battery V_{CB}. The directions of the emitter, base and collector currents are as shown in the figure. The direction of each current is opposite to the direction of motion of the electrons. The electrons being majority carriers in the emitter are repelled due to forward bias towards the base. The base contains holes as majority carriers and some holes and electrons combine in the base region. Since the base is lightly doped, the probability of electron-hole combination in base region is very small (5%). The remaining electrons cross into collector region, and enter into the positive terminal of the battery V_{CB} connected to the collector. At the same time an electron enters the emitter from the negative pole of the emitter-base battery V_{EB}. Thus, in *n-p-n* transistors, the current is carried inside the transistor as well as in the external circuit by the electrons. If I_E, I_B and I_C are respectively the emitter current, the base current and the collector current, then $I_E = I_B + I_C$.

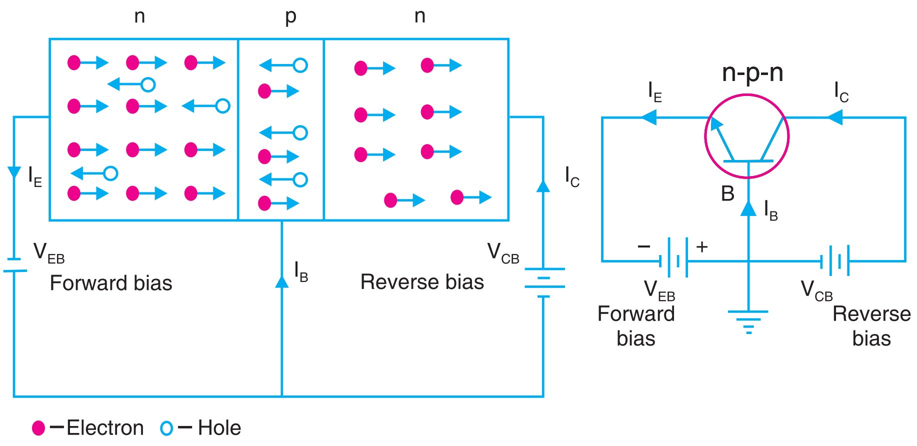


Fig. 58.23

58.15 Common Emitter Characteristics of a Transistor

Fig. 58.24 shows the circuit for drawing the common emitter characteristics of *n-p-n* transistor. The base-emitter junction is forward-biased by using base-emitter battery V_{BB}. The base-emitter voltage V_{BE} is read by a voltmeter and the base current I_B by a microammeter (μA). The base-collector junction is reverse biased by using collector-emitter battery V_{CC}. The collector-to-emitter voltage V_{CE} is read by another voltmeter and the collector current I_C by a milliammeter (mA).

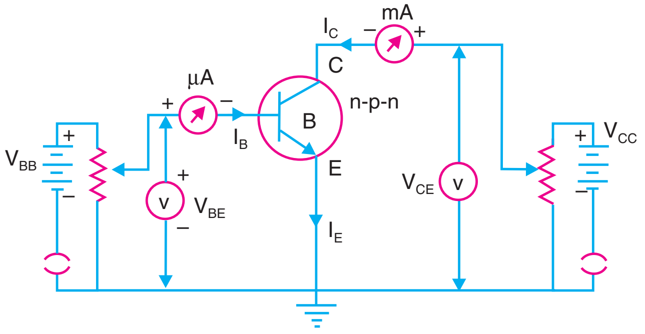


Fig. 58.24

1. Input characteristics. The curves showing the variation of the base current I_B with base-emitter voltage V_{BE} at constant collector-emitter voltage V_{CE} are called the input characteristics of a transistor.

To draw the input characteristics for *n-p-n* transistor, the collector-emitter voltage V_{CE} is maintained at a constant value, say 5V. The base currents I_B are noted for various values of input potential V_{BE}. A graph is drawn between V_{BE} taken along the *X*-axis and I_B taken along *Y*-axis. The experiment is repeated for different values of V_{CE}. Typical input characteristics are shown in Fig. 58.25 for $V_{CE} = 5$ and 20 volts. The shape of the graph is the same as that of the forward characteristics of *p-n* junction diode.

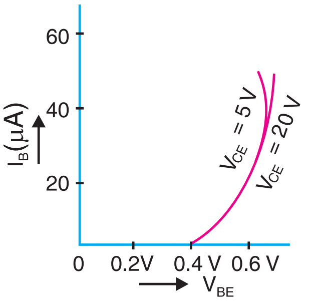


Fig. 58.25

$$\text{Input resistance, } r_i = \left(\frac{\Delta V_{BE}}{\Delta I_B}\right)_{V_{CE}}.$$

2. Output characteristics. Curves showing the variation of the collector current (I_C) with the collector-emitter voltage V_{CE} at constant base current I_B are called the output characteristics of a transistor.

The collector current I_C is measured by varying the collector voltage V_{CE} for a constant value of base current I_B. The experiment is repeated for different values of I_B and a family of curves is obtained as shown in Fig. 58.26.

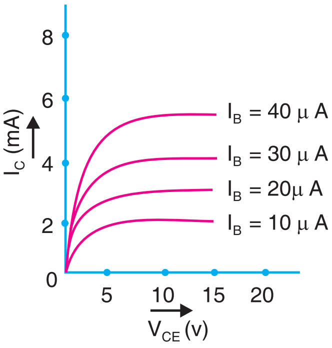


Fig. 58.26

A study of these curves reveals the following facts :

(*i*) The collector current I_C varies rapidly with V_{CE} for very small voltages (say up to $V_{CE} = 2$ volts). After this collector current becomes almost constant and is decided entirely by base current I_B. It then becomes independent of V_{CE}.

(*ii*) The collector current I_C is not zero when base current is zero. This is due to the inherent intrinsic conduction in semiconductors.

(*iii*) **Current gain β :** This is defined as the ratio of change in collector current to the change in base current for constant value of collector voltage V_{CE} in common emitter configuration, *i.e.*,

$$\beta = \left(\frac{\Delta I_C}{\Delta I_B}\right)_{V_{CE} = \text{constant}}.$$

β ranges from 20 to 200.

Since the input current I_B is measured in microampere and output current I_C is measured in milliampere, the common emitter configuration exhibits a current amplification β. Thus small input current I_B produces a large output current I_C or in other words, the collector current is β times greater than the base current.

(*iv*) Output resistance $$r_o = \left(\frac{\delta V_{CE}}{\delta I_C}\right)_{I_B = \text{constant}}.$$

58.16 Common-Emitter Transistor Amplifier

Fig. 58.27 shows the common-emitter amplifier circuit using *n-p-n* transistor. The emitter is common to both the input and output circuits. The emitter is forward biased by using base-emitter battery V_{BB}. Due to the forward bias, the resistance of input circuit is low. The collector is reverse biased by using collector-emitter battery V_{CC}. The low input voltage signal is applied in base-emitter circuit (input circuit). The amplified output is obtained across the collector and emitter.

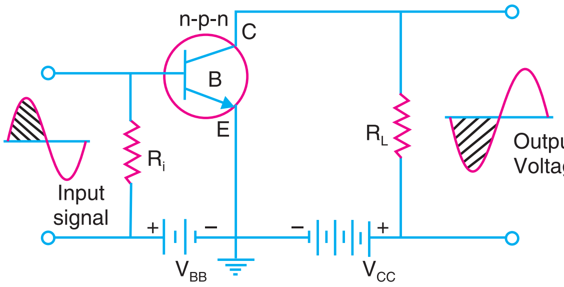


Fig. 58.27

In a common-emitter circuit, the collector-current is controlled by the base-current rather than the emitter-current. We know that in a transistor a large collector-current corresponds to a very small base-current. Therefore, with input signal applied to the base, a very small change in base-current produces a much more larger change in the collector-current. Thus, a very substantial *current-gain* is obtained.

(*a*) D.C. current gain : It is the ratio of the collector current to the base current. It is denoted by β.

$$\therefore \qquad \text{d.c. current gain, } \beta = \frac{I_C}{I_B}.$$

(*b*) A.C. current gain : It is defined as the ratio of the change in collector current to the change in base current at constant collector voltage. It is denoted by β_{ac}.

$$\therefore \qquad \beta_{ac} = \left(\frac{\Delta I_C}{\Delta I_B}\right)_{V_{CC}}$$

Its value is quite large as compared to 1 and lies between 150 and 50.

(*c*) A.C. voltage gain : It is the ratio of the change in output voltage to the change in input voltage. It is denoted by A_V.

$$A_V = \beta_{ac} \times \text{resistance gain} .$$

(*d*) A.C. power gain $= \dfrac{\text{Change in output power}}{\text{Change in input power}}$

$$\text{a.c. power gain } = \beta^2_{a.c.} \times \text{resistance gain} .$$

(*e*) Phase relationship. In the common-emitter amplifier the output signal is in antiphase with the input signal.

58.17 Practical Circuit of Common-Emitter Amplifier

A transistor gives a faithful amplification only if proper associated circuit elements are used. Fig. 58.28 shows a practical single stage transistor amplifier. Its various circuit elements and their functions are explained below :

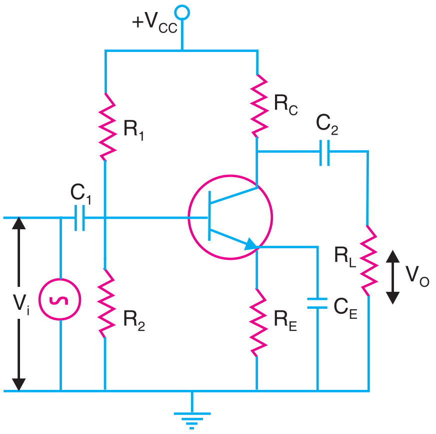


Fig. 58.28

(*i*) Biasing circuit. The resistors R_1, R_2, and R_E form the biasing and stabilisation circuit. The biasing circuit must establish a proper operating point. Otherwise a part of the negative half cycle of signal may be cut off in the output.

(*ii*) Input capacitor C_1. An electrolytic capacitor (C_1) is used to couple the a.c. signal voltage to the base of the transistor. If it is not used, the resistance of the signal source will be in parallel to R_2 and the bias voltage of the base will change. Thus C_1 allows only a.c. from the signal source to flow into the input circuit and isolates the signal source from R_2.

(*iii*) Emitter bypass capacitor C_E. An emitter bypass electrolytic capacitor is used in parallel with R_E to provide a low resistance path to the amplified A.C. signal. If it is not used, then amplified A.C. signal flowing through R_E will cause a voltage drop across it, thereby reducing the output voltage.

(*iv*) Coupling capacitor C_2. The coupling electrolytic capacitor C_2 couples one stage of amplification to the next stage. It prevents the d.c. component of the output of the first stage from reaching the input of the second stage. But it allows the passage of A.C. signal. If it is not used, R_C will come in parallel with the resistance R_1 of the biasing network of the next stage and the bias conditions of the next stage will be drastically changed.

Working. When a weak A.C. signal is applied to the base of transistor, a small base current starts flowing. Due to the transistor action a much larger A.C. current flows through the collector load R_C. As the value of R_C is quite high, a large voltage drop appears across R_C. Thus a weak signal appears in amplified form in the collector-circuit. In this way, transistor acts as an amplifier.

58.18 A.C. Equivalent Circuit

Both d.c. and a.c. conditions prevail in a transistor amplifier. When no signal is applied, d.c. base current I_B, collector current I_C and emitter current I_E flow due to biasing circuit. With the application of signal, a.c. base current i_b, a.c. collector current i_c and a.c. emitter current i_e also flow. However, the total current in any branch is the sum of d.c. and a.c. currents through the branch.

For drawing a.c. equivalent circuit of transistor amplifier, we consider only the a.c. conditions. Thus,

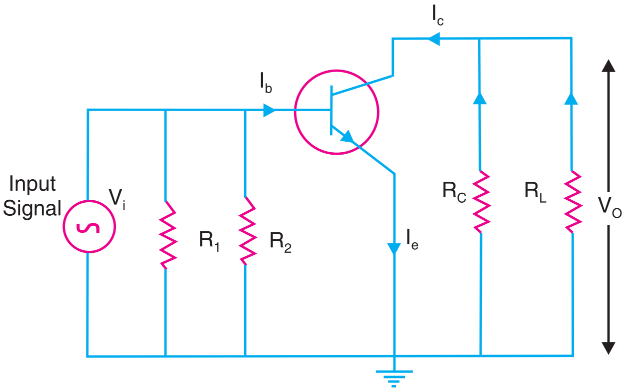


Fig. 58.29

(*i*) All d.c. sources are reduced to zero (*i.e.*, $V_{CC} = 0$).

(*ii*) All the capacitors are short circuited.

The a.c. equivalent circuit is shown in Fig. 58.29. It is used to calculate the voltage gain.

Collector load R_C appears in parallel with load R_L. Therefore, effective load R_{AC} for a.c. is given by

$$R_{AC} = \frac{R_C\, R_L}{R_C + R_L}$$

Output voltage, $V_0 = i_c\, R_{AC}$

Input voltage, $V_i = i_b\, R_i$

where R_i is the input impedance.

Voltage gain (A_v) is the ratio of a.c. output voltage to the a.c. input signal voltage.

$$\therefore \quad A_V = \frac{V_0}{V_i} = \frac{i_c\, R_{AC}}{i_b\, R_i} = \beta\, \frac{R_{AC}}{R_i}$$

where $\beta = \dfrac{i_c}{i_b}$ is the alternating current amplification factor.

Power gain. Power gain is given by

$$\text{Power gain} = \text{current gain} \times \text{voltage gain}$$

$$A_p = \frac{i_c}{i_b} \times \frac{i_c\, R_{AC}}{i_b\, R_i} = \beta^2 \times \frac{R_{AC}}{R_i}\,.$$

58.19 Common-Base Transistor Amplifier

Fig. 58.30 shows the common-base amplifier circuit using *n-p-n* transistor. The weak signal to be amplified is applied between the emitter and the base. The output is taken across the load R_L connected between the collector and the base. The d.c. voltage V_{EE} is applied in the input circuit in order to keep the input section always forward biased irrespective of polarity of the input signal. The d.c. voltage V_{CC} is chosen such that it keeps the output section always reverse biased.

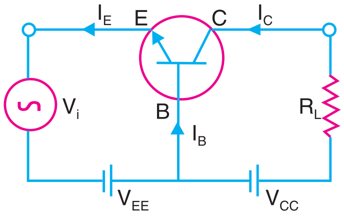


Fig. 58.30

The input circuit has a low resistance. Hence a small change in the signal voltage will cause an appreciable change in the emitter current. This causes almost the same change in the collector current due to transistor action. Thus, collector current flowing through the load resistor causes a large voltage drop. Thus a weak signal applied in the input circuit gets amplified in the collector circuit. *The output voltage is in phase with the input signal.*

Current amplification factor : The ratio of the change in the collector current to the change in the emitter current at constant collector-base voltage (V_{CB}) is called the current amplification factor α.

Thus $$\alpha = \left(\frac{\Delta I_C}{\Delta I_E}\right)_{V_{CB}}.$$

The current amplification factor is always less than 1.

Relation between α and β. By definition,

$$\alpha = \left(\frac{\Delta I_C}{\Delta I_E}\right)_{V_{CB}} \quad \text{and} \quad \beta = \left(\frac{\Delta I_C}{\Delta I_B}\right)_{V_{CE}}$$

For both the types of amplifier,

$$I_E = I_B + I_C.$$

$$\therefore \quad \Delta I_E = \Delta I_B + \Delta I_C$$

Dividing by ΔI_C, $$\frac{\Delta I_E}{\Delta I_C} = \frac{\Delta I_B}{\Delta I_C} + 1$$

$$\frac{1}{\alpha} = \frac{1}{\beta} + 1$$

$$\therefore \quad \beta = \frac{\alpha}{1-\alpha}.$$

Example 1. *In a transistor circuit emitter-base junction is usually forward biased while the collector-base junction is reverse biased. Justify.*

Solution. For a transistor to act as an amplifier, the emitter-base junction (input section) should offer a low resistance path for the signal to be amplified. Hence it is forward biased. As the output is realised across the collector-base junction, there should be large voltage drop across the load. Hence the collector-base junction is always reverse biased.

Example 2. *In a transistor the emitter and collector are of the same type of semiconducting material. Yet they cannot be interchanged in a circuit connection. Explain.*

Solution. Collector is wider than the emitter. Emitter is heavily doped compared to collector. Thus, they are physically different. The function of the emitter is to emit the majority carriers while the collector collects it. Hence, they are functionally different. That is why they cannot be interchanged.

Example 3. *A transistor is connected in the CE configuration. The collector supply voltage is* 10 *V and the voltage drop across the* 500 Ω *connected in the collector circuit is* 0.6 *V*. If $\alpha = 0.96$, *find the* (a) *collector-emitter voltage,* (b) *base current, and* (c) *the emitter current.*

Solution. The collector current, $I_C = \dfrac{V_C}{R_C} = \dfrac{0.6}{500} = 1.2$ mA

(1) Collector-emitter voltage $V_{CE} = V_{CC} - V_C = 10 - 0.6 = 9.4$ V

(2) $$\alpha = \frac{I_C}{I_E}$$

$$\therefore \quad I_E = \frac{I_C}{\alpha} = \frac{1.2\,\text{mA}}{0.96} = 1.25 \text{ mA}$$

(3) $$I_E = I_B + I_C$$

$$\therefore \quad I_B = I_E - I_C = (1.25 - 1.2) \text{ mA} = 0.05 \text{ mA.}$$

Example 4. *The constant* α *of a transistor is* 0.95. *What would be the change in the collector-current corresponding to a change of* 0.4 *mA in the base-current in a common-emitter arrangement*?

Solution. The current-gain of a transistor in common-emitter arrangement is β, which is related to its current-gain α in common-base arrangement by

$$\beta = \frac{\alpha}{1-\alpha}.$$

Here $\alpha = 0.95.$

$$\therefore \quad \beta = \frac{0.95}{1 - 0.95} = 19.$$

Now, β is the ratio of change in collector-current to the change in base-current.

$$\beta = \frac{\Delta i_c}{\Delta i_b}$$

$$\therefore \quad \Delta i_c = \beta \times \Delta i_b = 19 \times 0.4 \text{ mA} = \mathbf{7.6\ mA.}$$

Example 5. *Explain why CE configuration is widely used in amplifier circuits.*

Solution. A good amplifier stage is one which has a high input resistance and a low output resistance. A transistor in the *CB* configuration has a low input resistance (about 20 Ω) and a high output resistance (about 1 MΩ). This is just the opposite to what is desirable in a good amplifier. Hence the *CB* configuration is not useful for amplification. In a *CE* configuration, the input resistance is about 1 kΩ and the output resistance is about 10 kΩ. Therefore a transistor in the *CE* configuration makes a good amplifier. Further the current gain, voltage gain and power gain are all much greater in the *CE* configuration than in the *CB* configuration.

OSCILLATORS

58.20 Introduction

An oscillator is an electronic device for generating alternating current of a desired frequency. An oscillator converts dc power from the dc source to ac power at the load. Oscillation is achieved through positive feedback in amplifiers.

The essential components of an *LC* oscillator are shown in Fig. 58.31.

(*i*) Tank circuit. It is a parallel combination of inductor *L* and capacitor *C*. The frequency of oscillations in the circuit depends upon the values of inductance and capacitance.

$$f = \frac{1}{2\pi}\sqrt{\left(\frac{1}{LC}\right)}.$$

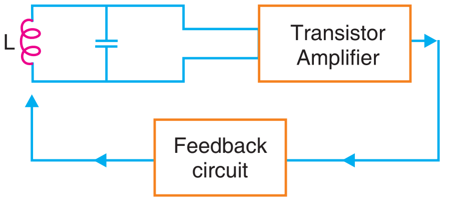
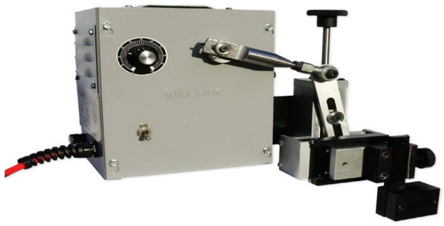


Fig. 58.31



Oscillators.

(*ii*) Transistor amplifier. The function of the amplifier is to amplify the oscillations produced by the tank circuit.

(*iii*) Feedback circuit. A positive feedback network to transfer a part of output energy to the tank circuit in proper phase. The amount of energy feedback is enough to compensate for the losses of energy in the *LC* circuit.

58.21 Condition for Self-Excitation (Barkhausen Criterion for Oscillations)

Consider a feedback amplifier (Fig. 58.32). Let A = voltage gain of amplifier without feedback, e_g = input voltage and e_0 = output voltage. Let the feedback network introduce a fraction β of the output into the input. If the feedback is positive, the voltage appearing at the input of the amplifier is $e_g + \beta e_0$. It is amplified A times by the amplifier to give the output e_0.

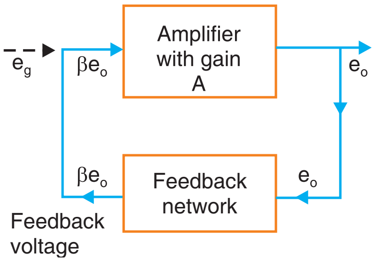


Fig. 58.32

$$\therefore \qquad (e_g + \beta e_0)\,A = e_0$$

$$\text{or} \qquad e_g = \frac{e_0\,(1 - \beta A)}{A}.$$

A is called open loop gain and $A\beta$ is called feedback factor or loop gain. If $\beta A = 1$, $e_g = 0$, *i.e.*, an output voltage is obtained without any input signal. Then the amplifier functions as an oscillator. Therefore, the condition for the maintenance of oscillations is

$$\beta A = 1 \qquad \text{...(1)}$$

$$\text{or} \qquad \beta = \frac{1}{A} \qquad \text{...(2)}$$

Eq. (1) is called 'Barkhausen criterion' for self-sustained oscillations.

In practical oscillators, $\beta A > 1$.

In general, the amplifier gain A and the feedback factor β will be complex. This means that, satisfying Eqn. (1) involves two conditions:

(*i*) The magnitude of the product $A\beta$ should be unity.

(*ii*) The product of the phase angles of the two should be zero or an integral multiple of 2π.

58.22 Hartley Oscillator : See section 66.1, Chapter 66.

58.23 Colpitt's Oscillator : See section 66.2, Chapter 66.

58.24 Tunnel Diode Operation:

(*a*) A *p*-*n* junction between two degenerate semiconductors is illustrated in terms of energy bands in Fig. 58.33 (*a*).

In this equilibrium condition, the Fermi level is constant throughout the junction. E_{FP} lies below the valence band edge on the p side. E_{Fn} is above the conduction band edge on the n side. Thus the bands must overlap on the energy scale in order for E_F to be constant.

(*b*) Suppose, now a small reverse bias is applied [Fig. 58.33 (*b*)]. It allows electron tunneling from the filled valence band states below E_{Fp} to the empty conduction band states above E_{Fn}. As the reverse bias is increased, E_{Fn} continues to move down the energy scale with respect to E_{Fp}, placing more filled states on the p side opposite empty states on the n side.Thus the tunneling of electrons from p to n increases with increasing reverse bias. The resulting conventional current is opposite to the electron flow, that is, from n to p.

(*c*) When a small forward bias V is applied, E_{Fn} moves up in energy with respect to E_{Fp} by the amount eV. Thus electrons below E_{Fn} on the n side are placed opposite empty states above E_{Fp} on the p side [Fig. 58.33 (*c*)]. Electron tunneling occurs from n to p with the resulting conventional current from p to n. This forward tunneling current continues to increase with increased bias as more filled states are placed opposite empty states.

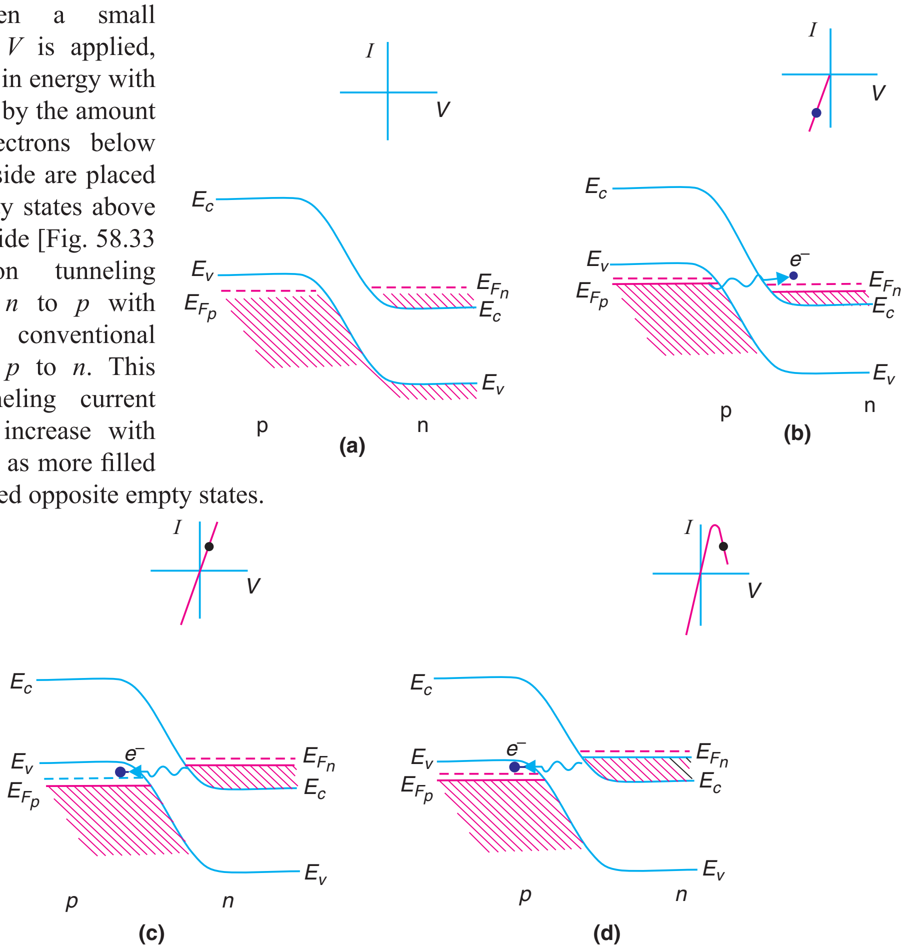


Fig. 58.33

The tunnel diode energy band diagrams: (*a*) zero-bias condition (no tunneling), (*b*) small reverse bias (electron tunneling from p to n), (*c*) small forward bias (electron tunneling from n to p), (*d*) increased forward biased (e^- tunneling from n to p decreases as bands pass by each other).

(*d*) However, as E_{Fn} continues to move up with respect to E_{Fp}, a point is reached at which the bands begin to pass by each other. When this occurs, the number of filled states opposite empty states decreases.The resulting decrease in tunneling current is illustrated in Fig. 58.33 (*d*). This region of the I–V characteristic corresponds to the *decrease* of tunneling current with *increased* bias, *i.e.*, it produces a region of negative slope. The *dynamic resistance* dV/dI is negative.

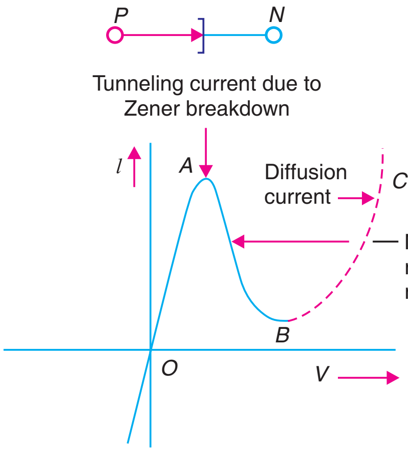


Fig. 58.34

Tunnel diode symbol and characteristic

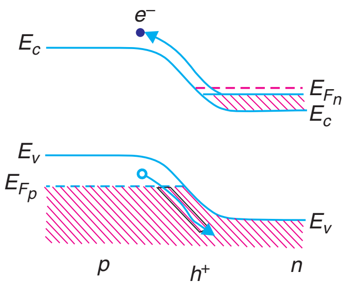


Fig. 58.35

Energy band diagram for the diode beyond the tunneling condition (corresponding to part *BC* of the curve in Fig. 58.34)

If the forward bias is increased beyond the negative resistance region, the current begins to increase again (Fig. 58.34). Once the bands have passed each other, the characteristic resembles that of a conventional diode. The forward current is now dominated by the diffusion current.

Fig. 58.35 shows the band diagram for the tunnel diode beyond the tunnel current region.

SPECIAL DEVICES

58.25 Tunnel Diode

It is a *p-n* junction semiconductor diode in which the concentration of impurity atoms is very large in *p* and *n* regions ($\approx 10^{24}/m^3$). The width of the depletion region is very small ($\approx 10^{-8}$ m). Since the depletion region is very narrow, electrons are capable of tunneling through from one side of the junction to the other at relatively low forward bias voltage. This phenomenon is called *tunneling*. This type of diode is called Tunnel Diode (Fig. 58.36)

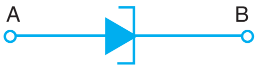


Fig. 58.36

MICROPHASE
CTM3818
110
120

Tunnel Diode

The Volt-Ampere (V–I) Characteristics

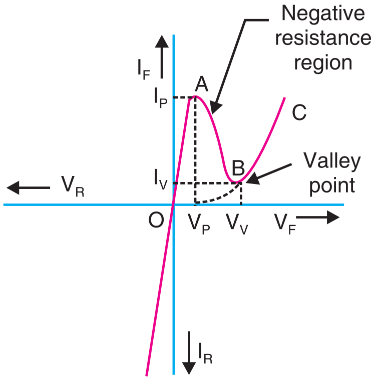


Fig. 58.37

The V–I characteristics of a tunnel diode are shown in Fig. 58.37. When forward voltage increases from zero, the forward current quickly reaches the peak value I_p (point A) at a particular low forward voltage V_p. When the forward voltage exceeds the value V_p, the forward current decreases and reaches a minimum value called valley current I_v at valley voltage V_v (point B). The region between the peak current I_p and the valley current I_v is called *negative resistance region*. When the forward voltage is further increased beyond the valley point, the current increases again as in ordinary *p-n* junction diode.

Explanation. Forward current in a tunnel diode is the sum of two components due to two different mechanisms:

(*i*) Normal injection current : It arises from an external voltage which reduces the potential barrier across the depletion region and allows current to flow due to majority carriers in the conduction band. It is shown by *dashed curve.*

(*ii*) Tunnel current : In a tunnel diode due to heavy doping, the depletion region is very narrow. It results in a very high electric field across the junction and allows carriers in the valence energy band on one side of the junction to 'tunnel' through to the conduction energy band on the other side of the junction without overcoming the potential barrier. The increase in current from 0 to its peak value I_P (at A) at a forward voltage V_P is due to tunneling phenomenon. The normal injection current is negligible at this value of forward bias.

The application of the *forward bias beyond* V_P reduces the electric field strength because the depletion region becomes less well defined due to diffusion of carriers across the *p-n* junction. Consequently the *tunneling current decreases.* Thus *current decreases with increase in applied voltage* between the peak point A and valley point B. It causes a negative slope in the *V-I* curve and the tunnel diode *exhibits a negative resistance property.*

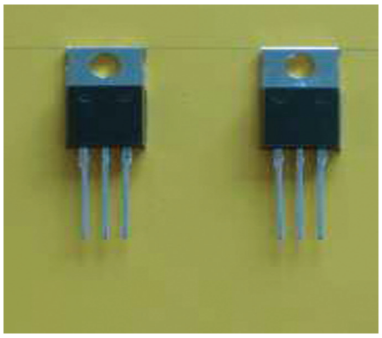
If forward voltage is increased beyond V_v, the tunneling effect ceases and current rises because of injection currents as in any ordinary junction diode.

Applications. The greatest advantage of the tunnel diode is its very fast response (change in current) to the voltage-change in the region *OAB*. In ordinary diodes and transistors, the response time depends on the diffusion speed of the charge-carriers which is low. Hence these devices cannot be used at high frequencies. Tunnel diodes, on the other hand, can be used as oscillators at frequencies above 10^{11} Hz, and in switching circuits operating in time-intervals ≈ 10^{-9} sec.

FIELD-EFFECT TRANSISTORS

58.26 Introduction

The field-effect transistor (FET) is a three terminal semiconductor device in which the output current is controlled by an applied electric field. Further the current in a FET is entirely due to the majority carriers whereas in a junction transistor both majority and minority carriers contribute to the current. Thus, while the junction transistor is bipolar, a FET is unipolar. There are two types of field-effect transistors : (*i*) *junction field effect transistor* (JFET or simply FET). (*ii*) *metal oxide semiconductor field effect transistor* (MOSFET).



Field-Effect Transistors.

A JFET can be either of the *n*-channel type or of the *p*-channel type. We shall here describe an *n*-channel JFET.

***n*-channel JFET.** It consists of a channel (*n*-type) into which two *p*-regions are diffused. One end of this symmetric structure is called *source* (*S*) and the other end is called *drain* (*D*). The two *p*-regions are connected together to a third terminal called *gate* (*G*) (Fig. 58.38). The *p*-regions are heavily doped compared to the *n*-region. During operation, majority carriers (electrons in this case) enter the channel through the source *S* and leave it through the drain *D*. The current is controlled by the gate which is always *reverse-biased.*

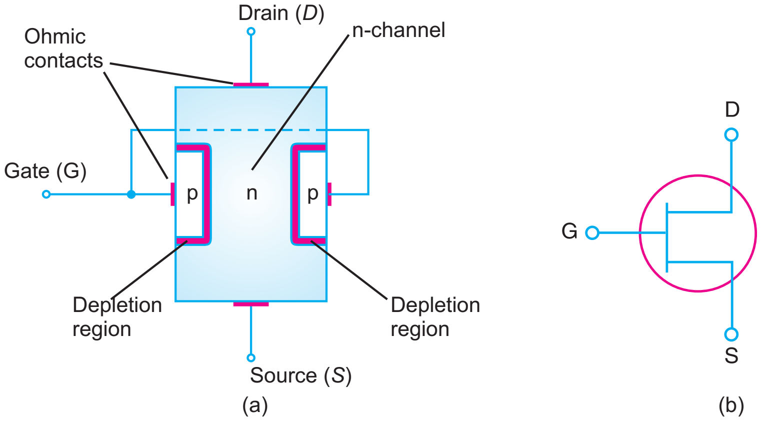


Fig. 58.38

58.27 Drain Characteristics of an n-Channel JFET

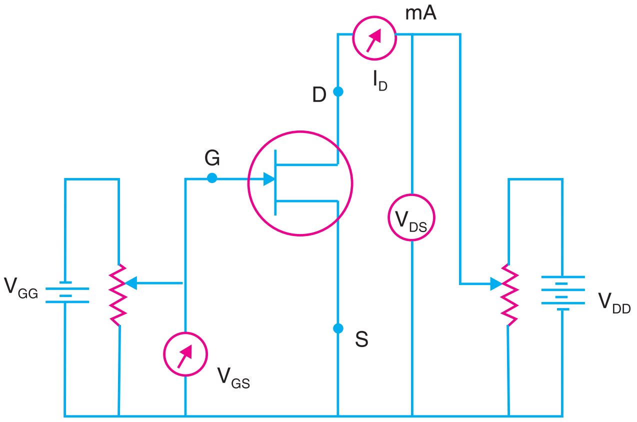


Fig. 58.39

A curve plotted between drain current i_D and drain-to-source voltage V_{DS}, at a fixed gate-to-source voltage V_{GS} is called the drain characteristics. Fig. 58.39 shows the circuit diagram for determining the output characteristics. V_{GG} is the gate bias supply and V_{DD} is the drain voltage source.

Keeping V_{GS} fixed at some value, the drain source voltage (V_{DS}) is changed in steps and the corresponding drain current I_D is noted. A group of such drain characteristics curves are drawn by setting V_{GS} at different fixed values. Fig. 58.40 shows a family of drain characteristics. There are three distinct regions in the characteristic thus obtained.

(*i*) When V_{DS} is small, the channel acts as a resistor. The current increases linearly with the voltage V_{DS} till point *A* is reached. This region of the characteristic is called the *ohmic region.*

(*ii*) When $V_{DS} = V_P$, the current I_D reaches its maximum value, I_{DSS}. If V_{DS} is increased beyond V_P, the current does not increase any further. The region *BC* is called saturation region or *pinch off region.*

(*iii*) *At a certain voltage* V_A, corresponding to point *C*, current increases suddenly due to avalanche breakdown. The covalent bonds in the depletion region break up and the current rises. This region is called the *avalanche region.* In actual practice this region is to be avoided.

Characteristic parameters of FET. There are three main characteristic parameters of a FET which describe its performance in an electronic circuit.

(*i*) Drain resistance r_d. It is defined as *the ratio of the change in drain-to-source voltage to the corresponding change in drain current at a constant gate-to-source voltage.*

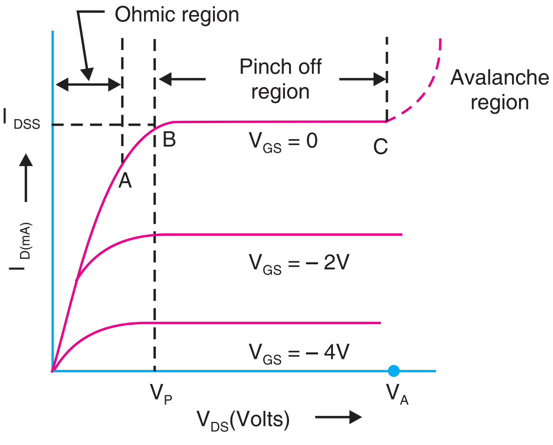


Fig. 58.40

$$r_d = \left(\frac{\Delta V_{DS}}{\Delta I_D}\right)_{V_{GS}} \qquad ...(i)$$

It is given by the reciprocal of the slope of the drain characteristic.

(*ii*) Transconductance g_m. It is defined as *the ratio of the change in drain current to the corresponding change in gate-to-source voltage at a constant drain-to-source voltage.*

$$g_m = \left(\frac{\Delta I_D}{\Delta V_{GS}}\right)_{V_{DS}} \qquad ...(ii)$$

The transconductance measures the control that the gate voltage has over the drain current.

(*iii*) Amplification factor μ. It is defined as *the ratio of the change in drain-to-source voltage to the corresponding change in gate-to-source voltage at a constant drain current.*

$$\mu = \left(\frac{\Delta V_{DS}}{\Delta V_{GS}}\right)_{I_D} . \qquad ..(iii)$$

Relation between the three parameters.

We have from the definition of the amplification factor

$$\mu = \frac{\Delta V_{DS}}{\Delta V_{GS}} = \frac{\Delta V_{DS}}{\Delta I_D} \times \frac{\Delta I_D}{\Delta V_{GS}} = r_d \times g_m$$

$$\therefore \quad \mu = r_d \times g_m.$$

(*ii*) Transfer characteristics

The graph drawn between drain current I_D and gate to source voltage V_{GS}, at constant V_{DS} is known as *Transfer characteristics* (Fig. 58.41).

For a constant value of V_{DS}, I_D is noted by varying V_{GS}. The value of V_{GS} (–ve) is varied till I_D becomes zero. This particular voltage V_P where $I_D = 0$ is called *pinch-off voltage.* The slope of the straight line portion of the graph gives the *transconductance* g_m.

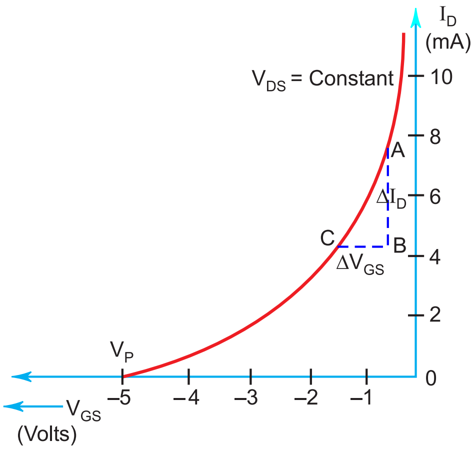


Fig. 58.41

Transconductance $$g_m = \left.\frac{\Delta I_D}{\Delta V_{GS}}\right|_{V_{DS}}$$

The transconductance is determined from the transfer characteristic curve.

$$g_m = \frac{AB}{BC} = \text{Slope of the curve.}$$

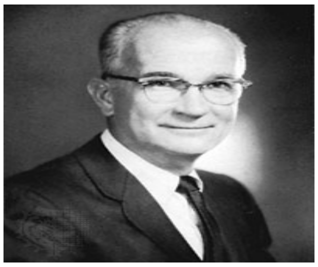
The relationship between I_D and V_{GS} is defined by *Shockley's equation.*

Control variable

$$\boxed{I_D = I_{DSS}\left(1 - \frac{V_{GS}}{V_P}\right)^2}$$

Constants

The squared term in the equation results in a nonlinear relationship between I_D and V_{GS}, producing a curve that grows exponentially with decreasing magnitude of V_{GS}.



William Bradford Shockley (1910-1989), co-inventor of the first transistor and formulator of the "field-effect" theory employed in the development of the transistor and the FET.

(Photo courtesy of AT & T Archives.)

58.28 Principle of n-Channel FET Operation

Case (*i*): When $V_{GS} = 0$ and $V_{DS} = 0$

When no voltages are applied between *D* and *S* and *G* and *S*, the thickness of the depletion regions around the *p-n* junctions is uniform. So a rectangular channel with uniform cross-section is formed [Fig. 58.38 (*a*)].

Case (*ii*): $V_{GS} = 0$ V, V_{DS} Some Positive Value

A positive voltage V_{DS} is applied across the channel and the gate is connected directly to the source to establish the condition $V_{GS} = 0$ V [Fig. 58.42 (*a*)].

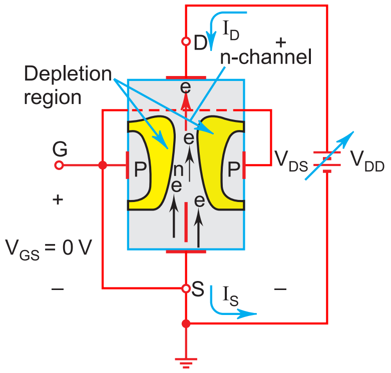


Fig. 58.42 (a)

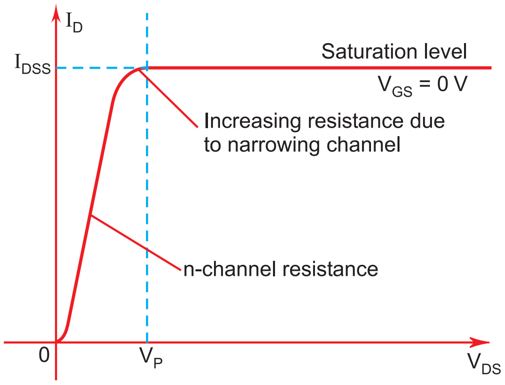


Fig. 58.42 (b)

The result is a gate and a source terminal at the same potential and a depletion region in the low end of each *p*-material similar to the distribution of the no-bias conditions of Fig. 58.38 (*a*). The instant the voltage V_{DD} (= V_{DS}) is applied, the electrons are drawn to the drain terminal, establishing the conventional current I_D with the defined direction of Fig. 58.42 (*a*). The path of charge flow reveals that the drain and source currents are equivalent ($I_D = I_S$). The flow of charge is relatively uninhibited and is limited solely by the resistance of the *n*-channel between drain and source. The depletion region is wider near the top of both *p*-type materials.

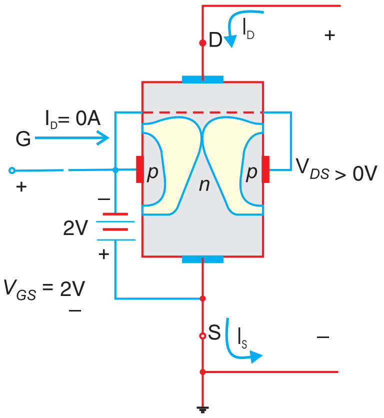


Fig. 58.42. (c)

Pinch-off (V_{GS} = 0 V, V_{DS} = V_P).

- As the voltage V_{DS} is increased from 0 V to a few volts, the current will increase as determined by Ohm's law. The plot of I_D versus V_{DS} will appear as shown in Fig. 58.42 (*b*). The relative straightness of the plot reveals that for the region of low values of V_{DS}, the resistance is essentially constant. As V_{DS} increases and approaches a level referred to as V_P, the depletion regions will widen, causing a noticeable reduction in the channel width. The reduced path of conduction causes the resistance to increase. The more horizontal the curve, the higher the resistance, suggesting that the resistance is approaching "infinite" ohms in the horizontal region. If V_{DS} is increased to a level where it appears that the two depletion regions would "touch" as shown in Fig. 58.42 (*c*), a condition referred to as *pinch-off* will result. The level of V_{DS} that establishes this condition is referred to as the *pinch-off voltage* and is denoted by V_P.
- As V_{DS} is increased beyond V_P, the region of close encounter between the two depletion regions increases in length along the channel, but the level of I_D remains essentially the same.
- *I_{DSS} is the maximum drain current for a JFET and is defined by the conditions*

$$V_{GS} = 0\text{ V} \quad \text{and} \quad V_{DS} > |V_P|.$$

Case (*iii*): $V_{GS} < 0$ V

The controlling voltage V_{GS} is made more and more negative from its V_{GS} = 0 V level. In other words, the gate terminal will be set at lower and lower potential levels as compared to the source. In Fig. 58.42 (*d*) a negative voltage of – 2V is applied between the gate and source terminals for a low level of V_{DS}. The effect of the applied negative-bias V_{GS} is to establish depletion regions similar to those obtained with V_{GS} = 0 V, but at lower levels of V_{DS}. Therefore, the result of applying a negative bias to the gate is to reach the saturation level at a lower level of V_{DS}, as shown in Fig. 58.40 for V_{GS} = – 2V. The resulting saturation level for I_D has been reduced and will continue to decrease as V_{GS} is made more and more negative.

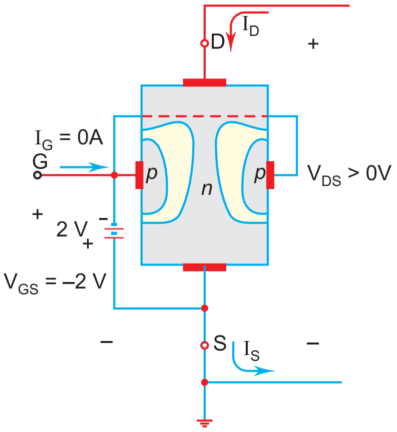


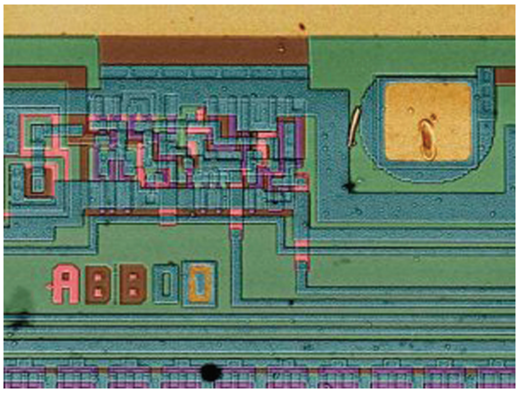
Fig. 58.42. (d)

INTEGRATED CIRCUITS

58.29 Integrated Circuit (IC)

An integrated circuit is just a package of electronic circuit in which both the active and passive components are fabricated on a small semiconductor chip.

Thus in an *IC*, a number of circuits containing many diodes, transistors, resistors, capacitors etc., are formed and connected within an extremely single tiny chip of semiconductor material. The typical size of an *IC* is 0.2 mm × 0.2 mm × 0.001 mm. The individual circuit components of an *IC* cannot be removed or replaced because each one of them is an integral part of the same semiconductor chip. The different components are isolated from each other by isolation diffusion within the crystal chip and are interconnected by an aluminium layer which serves as wires. Monolithic *IC* is the most commonly used one.



Integrated Circuit

58.30 Fabrication of a Monolithic IC

In monolithic *IC* the entire circuit (all circuit components and their inter-connections) is formed on a single thin wafer, called the substrate.

Planar technology for *IC* fabrication consists of the following processes:

1. Wafer (or substrate) preparation. A single and pure *p*-type silicon crystal is grown and cut into wafers which are cleaned and polished to a mirror finish. The typical thickness of a wafer is ≈ 200 μm. This wafer serves as the base for hundreds of ICs.

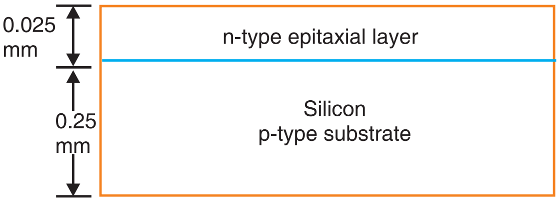


Fig. 58.43

2. Epitaxial growth. The epitaxial process is used to grow a thin layer (≈ 25 μm) of low resistivity *n*-type silicon on the high resistivity *p*-type substrate (Fig. 58.43). Epitaxial growth is performed in a special furnace called a *reactor* into which the finished silicon wafers are inserted and heated to 1200º C. Hydrogen reduction of $SiCl_4$ is used as the source of silicon to be grown. The chemical reaction is

$$SiCl_4 + 2H_2 \xrightleftharpoons{1200^\circ C} Si + 4HCl\,.$$

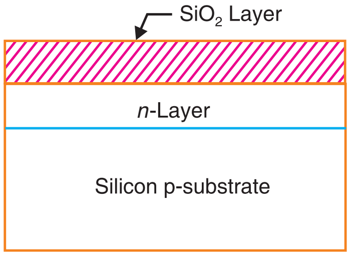


Fig. 58.44

For producing epitaxial film of particular impurity concentration, phosphine (PH_3) is added for *n*-type doping and biborane (B_2H_6) for *p*-type doping to the silicon tetrachloride-hydrogen gas stream.

The active and passive components are built within the thin *n*-type 'epitaxial layer' on top.

3. Oxidation. A thin (≈ 1 μm) layer of SiO_2 is formed over the entire wafer (Fig. 58.44) by heating the substrate to about 1000º C in an oxygen atmosphere. SiO_2 has the property of preventing the diffusion of impurities through it.

4. Photolithographic etching process. Here the SiO_2 layer is removed at different selected places on the surface of the wafer by a photolithographic etching process. It forms openings through which impurities may be diffused.

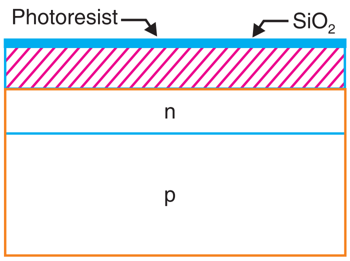


Fig. 58.45 (i)

Fig. 58.45 illustrates the photoetching method.

(*i*) The wafer is coated with a uniform film of a photosensitive emulsion called 'photoresist' [Fig. 58.45 (*i*)].

(*ii*) A large black-and-white layout of the desired pattern of openings is made and reduced photograpically. This negative is placed as a "mask" over the photoresist which is then exposed to ultraviolet light [Fig. 58.45 (*ii*)].

The photoresist under the transparent regions of the mask becomes polymerized. The mask is now removed, and the wafer is "developed" by a chemical (like trichloroethylene) which dissolves the unexposed (unpolymerized) portions of the photoresist coating and leaves the surface pattern as in Fig. 58.45 (*b*).

(*iii*) The SiO_2 not covered by polymerized photoresist is then removed by immersing the chip in an etching solution of hydro-fluoric acid. It does not affect the portions of SiO_2 which are protected by photoresist.

The exposed photoresist is then removed by a 'stripping process'. It leaves the wafer ready for diffusion process, as shown in Fig. 58.45 (*c*).

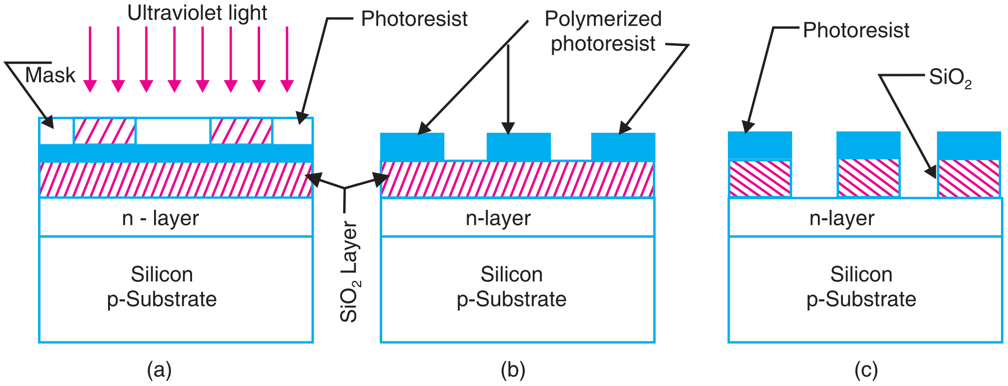

Ultraviolet light
Mask
Photoresist
Polymerized photoresist
Photoresist
SiO_2
n - layer
n-layer
n-layer
Silicon p-Substrate
Silicon p-Substrate
Silicon p-Substrate
SiO_2 Layer
(a)
(b)
(c)

Fig. 58.45 (ii)

5. Diffusion. The introduction of controlled impurity concentrations is performed in a diffusion furnace at a temperature of about 1000º C over a 1-to 2-*h* period. Gaseous impurities used are generally the hydrides of boron and phosphorus. An inert gas (nitrogen) transports the impurity atoms to the surface of the wafers where they diffuse into the silicon.

In a bipolar transistor, two diffusions of impurities are often used. For an *n-p-n* device, the first is the diffusion of the *p*-type base into the *n*-type epitaxially grown collector; the second is the *n*-type emitter region into the *p*-type base. Fig. 58.46 shows the cross-section of an IC transistor. The collector junction is formed at a distance of 2.7 μm from the surface. The emitter diffusion (phosphorus) starts from the surface and the emitter junction is formed at a depth of 2 μm. This junction corresponds to the intersection of the base and emitter distributions of impurities. We now see that the base thickness for this monolithic transistor is 0.7 μm.

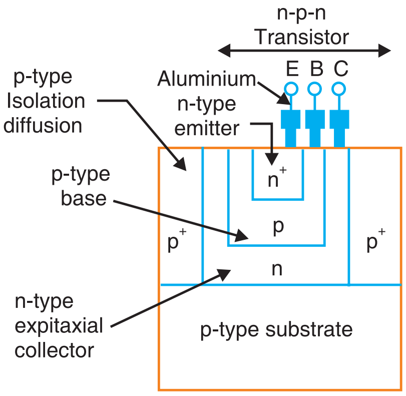

n-p-n Transistor
E B C
Aluminium
n-type emitter
p-type Isolation diffusion
p-type base
n^+
p
p^+
p^+
n
n-type expitaxial collector
p-type substrate

Fig. 58.46

6. Metallization. Interconnections between the various components of the integrated circuit are made by aluminium metallization. For this, a thin coating of aluminium is deposited over the entire wafer by vacuum evaporation of aluminium, and the undesired aluminium areas are then etched away. This leaves the desired pattern of ohmic contacts and interconnections, as shown in Fig. 58.46 in which an *n-p-n* transistor has been fabricated on a Si chip.

MODULATION AND DETECTION

58.31 Introduction

In radio broadcasting, an A.F. signal (speech, music, etc.) has to be transmitted to distant places. It is done by superimposing the given A.F. signal on an R.F. signal which will serve as a *carrier.* This process of superimposing an A.F. signal on an R.F. carrier wave is called *modulation.* The resulting wave is known as a modulated wave. At the receiving end, the original signal is extracted from the modulated wave. This is called *demodulation* or detection.

Modulation is defined as the process by which some characteristic property of a high frequency wave is varied in accordance with the instantaneous value of the signal.

The high frequency wave on which the signal is superimposed is called the carrier wave. Let the carrier wave be represented by

$$e_c = E_C \cos(\omega_c t + \theta).$$

Here, E_c, ω_c and θ are its peak amplitude, angular frequency, and phase respectively. In commercial transmitters, modulation is achieved by varying any one of the above three parameters in accordance with the signal, while keeping the remaining two parameters strictly constant. Thus, there are three types of modulation:

1. Amplitude modulation in which amplitude E_c is varied keeping ω_c and θ constant.
2. Frequency modulation in which the frequency ω_c is varied keeping E_c and θ constant.
3. Phase modulation in which the phase θ is varied keeping E_c and ω_c constant.

58.32 Amplitude Modulation

In amplitude modulation, the amplitude of the carrier is varied in accordance with the instantaneous values of the signal voltage. Here frequency and phase of the carrier remain constant. The waveforms for audio signal (modulating signal), carrier and amplitude modulated signal are shown in Fig. 58.47.

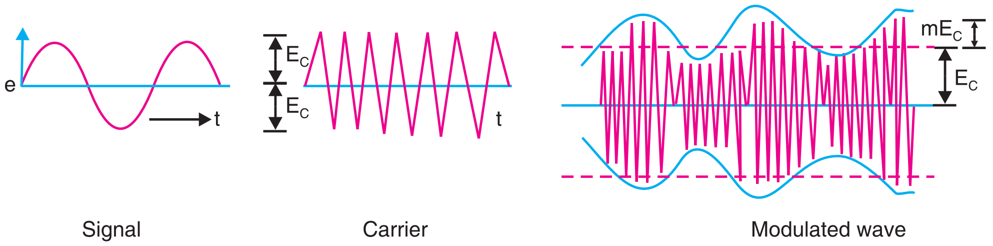


Fig. 58.47

Modulation index or modulation factor. It is defined as the ratio of the peak voltage (amplitude) of the signal to the peak voltage (amplitude) of the carrier, *i.e.*,

$$\text{Modulation Index } m = \frac{E_s}{E_c}.$$

Mathematical analysis of amplitude modulated wave

A carrier wave is represented by

$$e_c = E_c \cos \omega_c t \qquad \text{...(1)}$$

Here,
e_c = instantaneous carrier voltage,
E_c = amplitude of carrier,
$\omega_c = 2\pi f_c$ (f_c is frequency of carrier)

The audio signal (modulating signal) is represented by

$$e_s = E_s \cos \omega_s t \qquad \text{...(2)}$$

Here,
e_s = instantaneous voltage of signal
E_s = amplitude of audio signal.
$\omega_s = 2\pi f_s$, f_s = frequency of audio signal.

But, $E_s = mE_c$, where m is the modulation factor.

Therefore, the signal can be represented by

$$e_s = m E_c \cos \omega_s t \qquad \text{...(3)}$$

AM wave is obtained by varying E_c in accordance with e_s.

The amplitude of carrier wave varies at signal frequency f_s. Therefore, the amplitude of AM wave is given by

$$\begin{aligned} \text{amplitude} &= E_c + mE_c \cos \omega_s t \\ &= E_c (1 + m \cos \omega_s t) \end{aligned}$$

The instantaneous voltage of AM wave is

$$\begin{aligned} e &= \text{amplitude} \times \cos \omega_c t \\ &= E_c (1 + m \cos \omega_s t) \cos \omega_c t \\ &= E_c \cos \omega_c t + \frac{m E_c}{2}(2 \cos \omega_c t \cos \omega_s t) \\ &= E_c \cos \omega_c t + \frac{m E_c}{2}\left[\cos(\omega_c + \omega_s)t + \cos(\omega_c - \omega_s)t\right] \\ &= E_c \cos \omega_c t + \left[\frac{m E_c}{2}\right]\cos(\omega_c + \omega_s)t + \left[\frac{m E_c}{2}\right]\cos(\omega_c - \omega_s)t\,. \end{aligned}$$

1. The AM wave is equivalent to the summation of three sinusoidal waves.

(*a*) One having amplitude E_c and frequency f_c.
(*b*) The second having amplitude $m\,E_c/2$ and frequency $(f_c + f_s)$.
(*c*) The third having amplitude $m\,E_c/2$ and frequency $(f_c - f_s)$.

2. The AM wave contains three frequencies f_c, $f_c + f_s$ and $f_c - f_s$. The first frequency is the carrier frequency. Thus, the process of modulation does not change the original carrier frequency, but produces two new frequencies $(f_c + f_s)$ and $(f_c - f_s)$. They are called sideband frequencies.

3. The sum of carrier frequency and signal frequency, *i.e.*, $(f_c + f_s)$ is called upper sideband frequency. The difference between the carrier and the signal frequencies is called the lower sideband frequency, *i.e.*, $(f_c - f_s)$.

From the frequency spectrum (Fig. 58.48), it is clear that USF and LSF are symmetrical w.r.t. the carrier frequency.

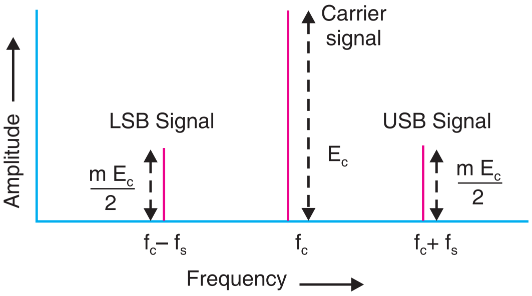


Fig. 58.48

Example 1. *In a broadcasting studio a 100 MHz carrier having an amplitude of 50 V is amplitude modulated by a 5 kHz audio signal having an amplitude of 20 V. Find the modulation index, percentage modulation, frequency components of the modulated wave and their amplitudes.*

Solution.

$$E_c = 50\text{V}, \quad E_s = 20\text{V}$$

$$E_{max} = E_c + E_s = 50 + 20 = 70 \text{ V}$$

$$E_{min} = E_c - \text{E}_s = 50 - 20 = 30 \text{ V}$$

The modulation index is given by,

$$m = \frac{E_{\max} - E_{\min}}{E_{\max} + E_{\min}} = \frac{70 - 30}{70 + 30} = 0.4$$

$$\text{Percentage of modulation} = 0.4 \times 100 = 40\%$$

$$\text{carrier frequency, } f_c = 100 \times 10^6 \text{ Hz}$$

$$\text{signal frequency, } f_s = 5 \times 10^3 \text{ Hz.}$$

The AM wave contains three frequencies, f_c, $f_c + f_s$, $f_c - f_s$

f_c –carrier frequency = 100 MHz

$f_c + f_s$ –upper sideband frequency

= 100 MHz + 5 kHz = 1,00,005 kHz

$f_c – f_s$ –lower sideband frequency

= 100 MHz –5 kHz = 99995 kHz

carrier amplitude = $E_c = 50$ V

$$\text{Amplitude of sidebands} = 2 \times m.\frac{E_c}{2} = 2 \times 0.4 \times \frac{50}{2} = 20 \text{ V}.$$

58.33 Power in an Amplitude Modulated Wave

The AM wave has three components:

(*a*) carrier wave of amplitude $E_c / \sqrt{2}$. (r.m.s. value)

(*b*) USF wave of amplitude $mE_c / \left(2\sqrt{2}\right)$. (r.m.s. value)

(*c*) LSF wave of amplitude $mE_c / \left(2\sqrt{2}\right)$. (r.m.s. value)

Total power P_T of the A.M. wave is the sum of the carrier power P_c and sideband power P_s.

$$P_T = P_c + P_s.$$

But power of a voltage wave is directly proportional to square of the amplitude.

$$\text{Carrier power} = P_c = KE_c^2 / 2$$ where K is a constant.

$$\text{Total sideband power} = P_s = 2 \times Km^2E_c^2 / 8 .$$

Total power carried by the AM wave

$$P_T = P_c + P_s = K\frac{E_c^2}{2}\left(1 + \frac{m^2}{2}\right) = P_c\left(\frac{2 + m^2}{2}\right).$$

$$\therefore \quad \text{carrier power} = P_c = P_T\left(\frac{2}{2 + m^2}\right)$$

$$\text{and total sideband power} = P_s = P_T\left(\frac{m^2}{2 + m^2}\right).$$

Example 1. *A carrier wave of* 500 *watt is subjected to* 100% *amplitude modulation. Determine* (1) *Power of modulated wave, and* (2) *Power in sidebands*.

Solution. Power of modulated waves is given by

$$P_T = P_c\,[1 + m^2/2]$$

Here, P_c = power in carrier wave = 500 watt

m = modulation index = 1 (for 100% modulation)

$$P_T = 500\,[1 + 1/2] = 750 \text{ watts}$$

Power in sidebands, $P_s = P_T - P_c = 750 - 500 = 250$ watt.

There is 125 watt power in upper sideband and 125 watt in lower sideband.

Example 2. *The total power content of an AM wave is* 1500 *watt for a hundred per cent modulation. Determine* (*i*) *power transmitted by carrier,* (*ii*) *power transmitted by each sideband. What is your inference*?

Solution. Carrier power $P_C = P_T\left(\dfrac{2}{2+m^2}\right)$

When modulation is 100%, then $m = 1$.

$\therefore$ Power transmitted by carrier $= 1500\left(\dfrac{2}{2+1}\right) = 1000$ W.

Power transmitted by sidebands $P_S = P_T\left(\dfrac{m^2}{2+m^2}\right)$

$= 1500\,(1/3) = 500$ W.

$\therefore$ Power transmitted by each sideband = 500/2 = 250 W.

The carrier power is that required for transmission and sideband power represents the signal content. Of these two, the power due to carrier is much greater than the other. But the message or signal is present only in the latter, *i.e.*, in the sidebands. Hence most of the power from the modulating source goes to a component which does not contain the signal. If we can suppress the unwanted, large power consuming carrier component, then the power drawn from the source will be less and whatever is drawn from the source will go to the useful sidebands. This type of transmission is economical and is called carrier suppressed sideband transmission.

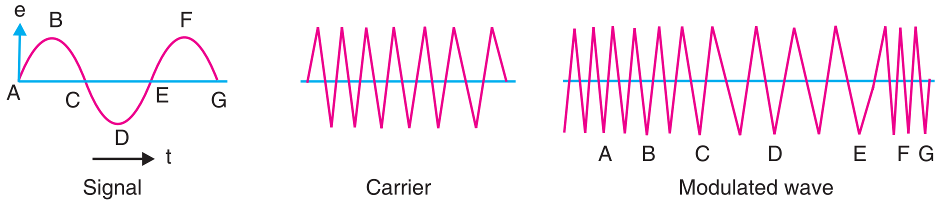
58.34 Frequency Modulation

In frequency modulation, the frequency of the carrier is varied in accordance with the instantaneous values of signal voltage. Here the amplitude and phase of the carrier remain constant.

The frequency of the carrier before modulation is called the centre frequency or resting frequency usually denoted by f_c. The variation of frequency either above or below the centre frequency is known as frequency deviation Δf. The frequency deviation is proportional to amplitude of the modulating signal. The maximum permitted frequency deviation is 75 kHz. So bandwidth for FM is 2 × 75 or 150 kHz. Bandwidth is also referred to as channel width.

In addition, a guardband of 25 kHz is allowed on either side of the channel width to eliminate the interference between the adjacent channels. Therefore in FM, total channel width is 200 kHz. But in AM the channel width is only 10 kHz.

Frequency modulation is illustrated in Fig. 58.49. The amplitude of the modulated wave remains the same as the carrier amplitude. The frequency of the carrier wave varies in proportion to the instantaneous value of amplitude of the signal. When the signal voltage is zero as at *A*, *C*, *E* and *G*, the carrier frequency is not changed. When the signal approaches its positive peaks as at *B* and *F*, the carrier frequency is increased to maximum as shown by the closely spaced cycles. However, during the negative peak of the signal, as at *D*, the carrier frequency is reduced to minimum as shown by the widely spaced cycles.


e
B
F
A
C
E
G
D
t
A
B
C
D
E
F
G

Signal Carrier Modulated wave

Fig. 58.49

58.35 Demodulation or Detection

Demodulation is the process of recovering the audio signal from the modulated wave. The process of detecting an AM carrier wave consists of two steps:

Step 1. The negative half-cycles of the modulated carrier wave have to be removed by rectification.

Step 2. The carrier frequency must be removed by using suitable filters. Fig. 58.50 illustrates the process of detecting an AM carrier wave.

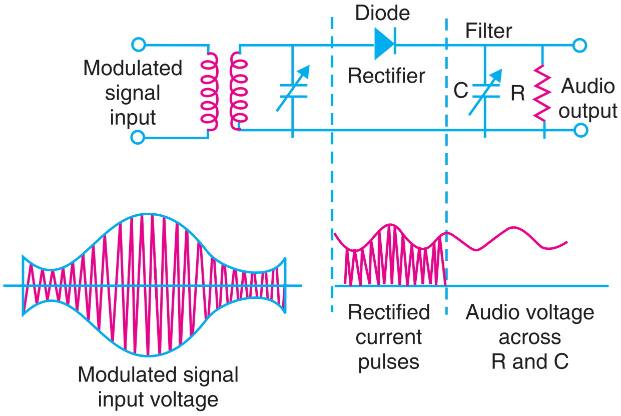

Diode
Filter
Modulated signal input
Rectifier
C
R
Audio output
Rectified current pulses
Audio voltage across R and C
Modulated signal input voltage

Fig. 58.50

58.36 Block Diagram of AM Transmitting System

Radio transmitters produce radio waves and transmit them. The information is contained in the audio signal. AM transmitter consists of (*i*) RF channel, and (*ii*) AF channel (Fig. 58.51).

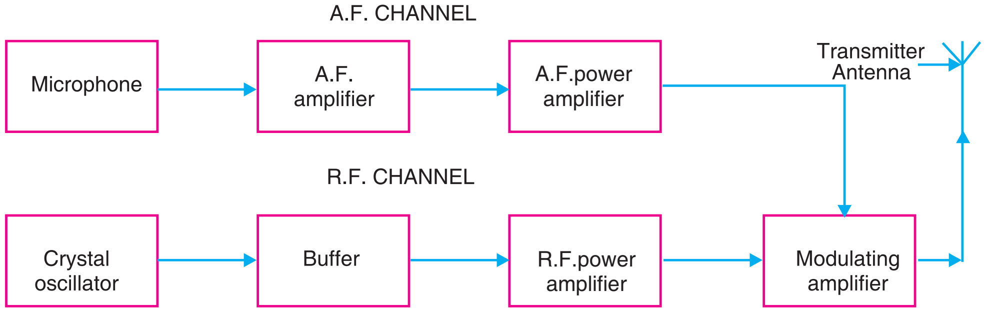

A.F. CHANNEL
Microphone
A.F. amplifier
A.F.power amplifier
Transmitter Antenna
R.F. CHANNEL
Crystal oscillator
Buffer
R.F.power amplifier
Modulating amplifier

Fig. 58.51

(*i*) RF Channel. The production, amplification and modulation of the carrier wave takes place in this part. The high frequency carrier wave is generated by a crystal controlled master oscillator. The output of the crystal controlled oscillator is power amplified by RF power amplifier. The buffer is a low gain amplifier which isolates the RF power amplifier from the oscillator. This arrangement keeps the frequency of the crystal controlled oscillator constant. The last stage of the RF channel is the modulating amplifier. In this amplifier, the RF wave and the modulating AF signal are mixed to produce the AM wave. Finally modulated waves are fed to transmitter antenna. The transmitter antenna radiates modulated waves into space.

(*ii*) AF Channel. Modulating signal is produced and amplified in this channel. The microphone converts sound waves into AF electrical signals. The AF signals are amplified first by an AF voltage amplifier and then by an AF power amplifier. The signal from this part is used as modulating signal. The output of the AF power amplifier is fed into the modulating amplifier.

58.37 Block Diagram of AM Receivers—Superheterodyne Receiver

In this receiver the incoming R.F. modulated signal, after amplification, is mixed with an unmodulated R.F. signal generated in the receiver itself by a local oscillator. This results in a new signal with the same modulation but a *lower* carrier frequency called 'intermediate frequency' (I.F.). This new signal is then amplified and detected. The intermediate frequency is usually 455 kHz.

Superheterodyne Receiver.

The block diagram of a superheterodyne receiver is shown in Fig. 58.52. The R.F. amplifier amplifies RF signal coming at the antenna. Its output is sent to the mixer where the RF signal f_s (say) is mixed with slightly lower or higher frequency f_0 from the local oscillator. When mixed, the sum $(f_0 + f_s)$ and difference $(f_0 - f_s)$ frequencies are formed. By a suitable arrangement only the difference frequency is extracted. The signal of intermediate frequency $(f_0 - f_s)$ is further amplified by the I.F. amplifier. The output of this stage is fed into the detector. Here, the A.F. modulating signal is separated from the carrier. The detected AF signal is amplified by the AF amplifier and then applied to a loudspeaker. The speaker converts audio electric signal into sound waves.

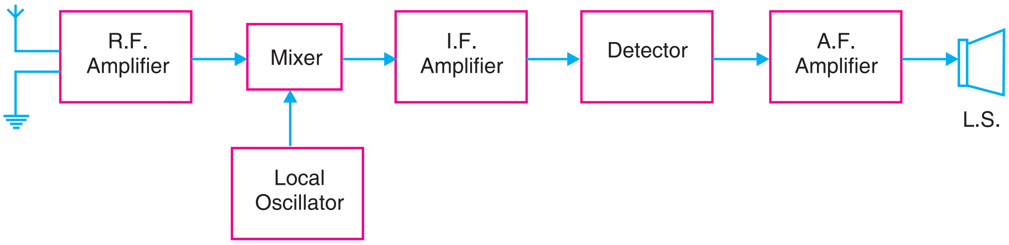


Fig. 58.52

58.38 Characteristics of a Radio-receiver

The performance of a radio-receiver is judged by its following characteristics :

(*i*) Sensitivity : It determines the minimum strength of an input signal necessary for producing a standard output signal.

(*ii*) Selectivity : It determines the extent to which the radio receiver can differentiate between the desired signal and the unwanted signals of other frequencies.

(*iii*) Fidelity : It is a measure of the receiver's ability to reproduce faithfully at its output all the essential characteristics of the input signal.

58.39 Block Diagram of TV Transmitting System

The block diagram of a typical TV transmitter is shown in Fig. 58.53. The picture (video) and sound (audio) signals are transmitted at different frequencies (to avoid interference) by a single antenna. These signals are processed separately at the TV station, as shown in the figure.

The video section consists of a TV camera tube which produces a video signal of the picture to be televised. The synchronising and scanning circuits provide the necessary sweep voltages to the vertical and horizontal deflector plates. Moreover, these synchronising pulses are needed for the reassembly of the image at the receiver in step with the scanning at the transmitter. So these signals are transmitted along with the video signals. The video signals are amplified in a video amplifier. The video signals are then used to amplitude modulate (AM) a radio frequency (RF) carrier in an AM video transmitter. The modulated wave is fed to the transmitting antenna.

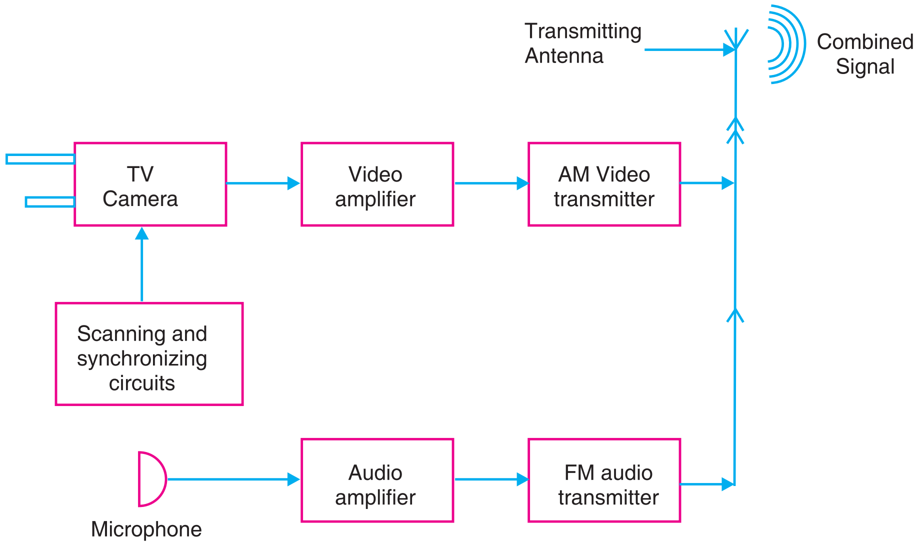


Fig. 58.53

In the audio section, the sound of the scene is picked up by a microphone. The microphone converts the audio signals into electrical signals. The electrical signals are amplified in an audio amplifier. Then the audio signals are used to frequency modulate a RF carrier whose frequency is 5.5 MHz greater than the video carrier. This frequency modulated (FM) sound (or audio) signal is also carried to the same transmitting antenna. The combined (AM video and FM audio) signal is radiated by the antenna in the form of electromagnetic waves.

58.40 Block Diagram of TV Receiving System

Fig. 58.54 shows a block diagram of TV receiver. It uses the superheterodyne principle. The incoming signal, consisting of audio and video RF signals, is picked up by the receiving antenna. The selected signals are amplified by the RF amplifier. The output of the RF amplifier is fed to the frequency mixer. The mixer heterodynes the amplified RF signal with the output of the local oscillator. It produces two intermediate frequencies, one for the picture and the other for the sound signal. Usually picture IF is 38.9 MHz and sound IF is 33.4 MHz. These frequencies are separated by the help of tuned circuits in the output of mixer stage. The output of the mixer is fed to two sets of tuned IF amplifier sections—one tuned to sound IF and the other to picture IF.

The sound IF signal thus obtained is applied to separate sound portion of the receiver. The amplified audio signal from IF amplifier is delivered to the *limiter stage* which *removes any interference riding on the top of the signals in the form of amplitude modulation.* The sound signal then passes through a discriminator (or a FM detector). The FM detector extracts audio signal from FM carrier. The audio signal is fed to an audio amplifier which amplifies the signal. The amplified signal is given to the loudspeaker which produces the original sound.

The video signal is amplified by a video amplifier. After detection and amplification, the video signals obtained are identical to those produced by the orthicon and the pulse generator at the transmitter. The output from the video amplifier is applied to the *control grid of the picture tube* or *kinescope.* When electrons strike the fluorescent screen of the kinescope, high energy luminous radiations come out. The brightness of these radiations varies with the variation of the amplitude of the picture signal and hence with the brightness of the transmitted image. Dark and bright spots thus appear on the screen in step with the dark and bright spots of the picture televised. These spots merge to give the exact replica of the picture televised.

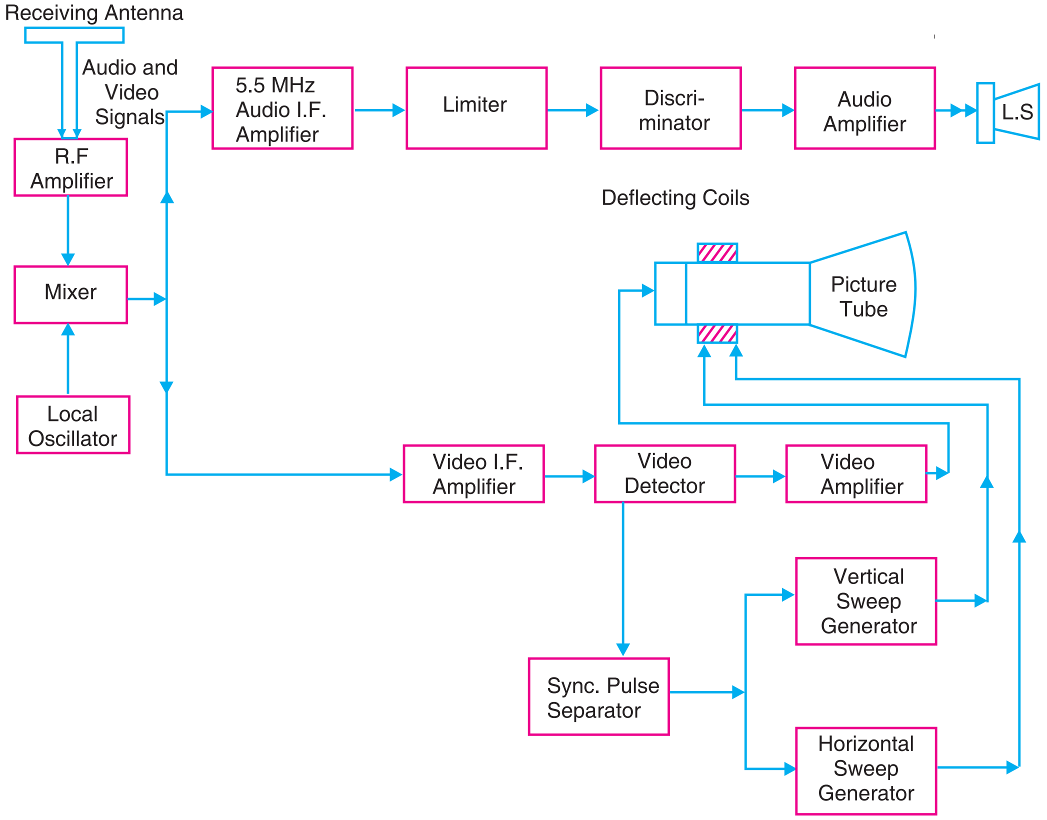

Receiving Antenna
Audio and Video Signals
R.F Amplifier
Mixer
Local Oscillator
5.5 MHz Audio I.F. Amplifier
Limiter
Discri-minator
Audio Amplifier
L.S
Deflecting Coils
Picture Tube
Video I.F. Amplifier
Video Detector
Video Amplifier
Vertical Sweep Generator
Sync. Pulse Separator
Horizontal Sweep Generator

Fig. 58.54

EXERCISE

1. Draw characteristics of a Zener diode. Explain with the help of a circuit diagram its use as voltage regulator. (*Rohilkhand,* 90)
2. Draw the circuit diagram and explain the working of a half wave diode rectifier. Explain ripple voltage and ripple factor. (*Lucknow* 80, *Rajasthan,* 84)
3. Explain giving circuit diagram the working of full wave rectifier. Obtain expressions for average and RMS values of current and efficiency. (*Rohilkhand,* 89)
4. Describe an experiment to study (*a*) the input; and (*b*) the transfer characteristics of an n-p-n transistor in the common emitter configuration. (*Calicut*, 1989)
5. How can the static characteristics of a field effect transistor be studied? Give the necessary circuit diagram. (*Kolkata,* 82)
6. Why is modulation employed in the transmission of information? Explain the salient features of amplitude modulation with the help of appropriate wave-forms. Explain how A.M. signals are detected using a diode. (*Kakatiya,* 93)

CHAPTER
59

TRANSISTOR BIASING

AT A GLANCE

59.1 Transistor Biasing
59.2 Fixed Bias or Base Bias
59.3 Collector to Base Bias Circuit
59.4 Self Bias or Emitter Bias (Voltage-Divider Bias)
59.5 Thermal Runaway
59.6 Characteristics of a Transistor Connected in Common Base Configuration
59.7 C – C Configuration

59.1 Transistor Biasing

A transistor is used as an amplifier. For faithful amplification, the following three conditions must be satisfied :

(*i*) The emitter-base junction should be forward-biased,
(*ii*) The collector-base junction should be reverse-biased, and
(*iii*) There should be proper zero signal collector current.

The proper flow of zero signal collector current (*proper operating point of a transistor*) *and the maintenance of proper collector-emitter voltage during the passage of signal is known as transistor biasing.*

When a transistor is not properly biased, it works inefficiently and produces distortion in the output signal.

Requirements of biasing circuit

Following are the requirements of a proper biasing circuit :

(*i*) Proper selection of operating point (*Q*-point) such that on applying the input signal, the instantaneous operating point does not move either to the saturation region or to the cut-off region.

(*ii*) To stabilize the collector current against temperature variations.

(*iii*) To make the operating point independent of transistor parameters (*i.e.*, if the transistor is replaced by another transistor of the same type, the operating point should not shift).

Stability Factor. The transistor parameters β, I_{C0} and V_{BE} are functions of temperature. Among these, the change in I_{C0} with temperature is more significant than other changes.

The stability factor S is defined as the rate of change of collector current I_C with respect to the reverse saturation current I_{C0} keeping β and V_{BE} constant, *i.e.*,

$$S = \frac{\partial I_C}{\partial I_{C0}} \approx \frac{\Delta I_C}{\Delta I_{C0}}.$$

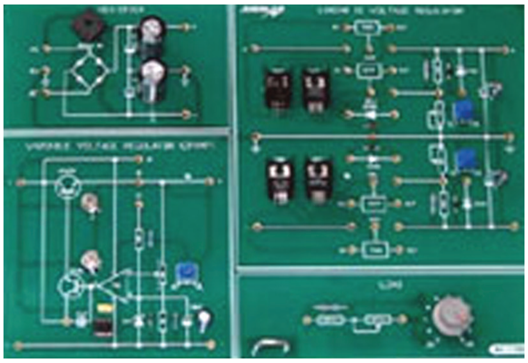
This expression shows that smaller is the value of S, higher is the stability. So the stability factor S should be kept as small as possible. The lowest value of S, that can be obtained, is unity since I_C must include I_{C0}. Closer is value of S to unity, lesser will be the variation of operating point with temperature.

There are three types of biasing techniques.

59.2 Fixed Bias or Base Bias

In this method, a high resistance R_B is connected between the base and the positive end of the power supply V_{CC}, for *n-p-n* transistor (Fig. 59.1). Let us determine the value of R_B, so that required collector current I_C (= β I_B) flows in the zero signal conditions.

Applying Kirchhoff's voltage law to the base-emitter loop *ABENA*, we get



Base Bias

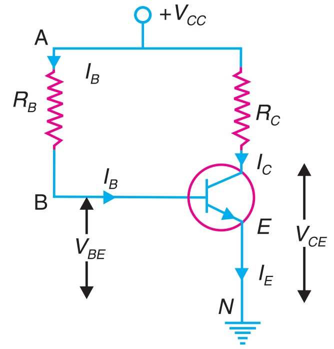


Fig. 59.1

$$V_{CC} = I_B R_B + V_{BE}$$

or

$$I_B R_B = V_{CC} - V_{BE}$$

$$\therefore \quad R_B = \frac{V_{CC} - V_{BE}}{I_B}$$

The value of V_{BE} is negligibly small, when compared with V_{CC}.

$$\therefore \quad R_B = \frac{V_{CC}}{I_B}$$

The supply voltage V_{CC} is fixed. Once the value of R_B is selected, the base current I_B is also fixed. Hence the name fixed bias circuit.

Advantages

1. The circuit is very simple since only one resistor is required to bias the base.
2. The biasing conditions can be easily set up.
3. There is no loading of the source by the biasing circuit.

Disadvantages

1. Stabilisation is poor. The collector current $I_C = \beta I_B$. Variation of β will cause corresponding variation of collector current. Such collector current variation cannot be checked by the fixed bias circuit. The operating point is highly unstable.
2. The possibility for thermal runaway is large.

This type is used rarely.

59.3 Collector to Base Bias Circuit

A resistance R_B is connected between the collector and base (Fig. 59.2). The bias voltage is obtained from the collector of the transistor. Here the required zero signal base current is determined by the collector-base voltage V_{CB} and not by V_{CC}. The required value of R_B needed to give the zero signal collector current I_C can be determined by applying Kirchhoff's voltage law to Fig. 59.2.

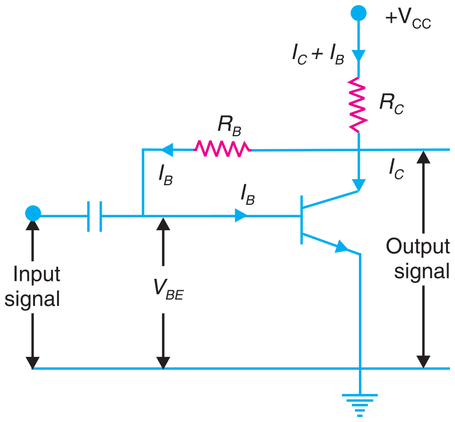


Fig. 59.2

$$V_{CC} = (I_C + I_B)\,R_C + I_B\,R_B + V_{BE}$$

$$V_{CC} = (I_C\,R_C + I_B\,R_B + V_{BE}) \qquad (\text{since } I_B << I_C)$$

or $$V_{CC} = \beta I_B\,R_C + I_B\,R_B + V_{BE} \qquad (\because\ I_C = \beta\,I_B)$$

or $$R_B = \frac{V_{CC} - V_{BE} - \beta\,I_B\,R_C}{I_B}\,.$$

Stabilisation. This circuit provides stabilisation of the operating point as discussed below.

When I_C tends to increase either because of rise in temperature or because of the transistor being replaced by another of larger β, then voltage V_{CE} decreases due to greater drop across R_C (since $V_{CE} = V_{CC} - I_C\,R_C$). Consequently, lesser voltage is available across R_B. Hence base current I_B decreases which in turn tends to decrease the collector current to its original value. Thus a mechanism exists in the circuit which does not allow the collector current to increase and hence the operating point is stabilised.

Disadvantages. R_B not only provides a dc feedback for the stabilisation of operating point but it also causes a feedback of a.c. signal from output (collector) terminal to input (base) terminal. This reduces the voltage gain of the amplifier.

59.4 Self Bias or Emitter Bias (Voltage-Divider Bias)

The circuit arrangement is shown in Fig. 59.3. Two resistors R_1 and R_2 are connected across supply voltage V_{CC} and provide biasing. The emitter resistor R_E provides stabilisation. R_1 and R_2 form a potential divider across V_{CC}. The resistor R_E causes a voltage drop in a direction so as to reverse-bias the emitter junction. Since the junction must be forward biased, the base voltage is obtained from the supply through $R_1 - R_2$ network. The net forward bias across the emitter junction is equal to V_B minus the d.c voltage drop across R_E.

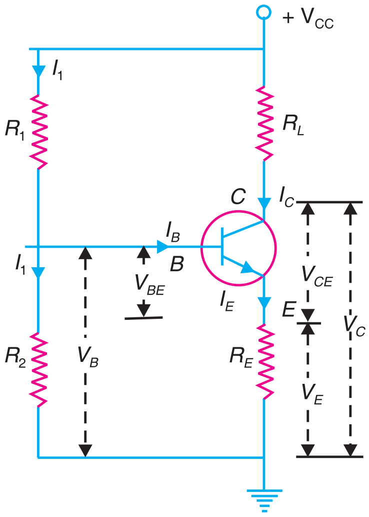


Fig. 59.3

The improvement in the operating point stability may be explained as follows : Let there be a rise in temperature. This causes a rise in I_{C0}, a rise in I_C, and in I_E. Now the current in R_E increases. As a result, the voltage drop across R_E increases and consequently the base current decreases. This decreases the collector current. Thus the presence of R_E reduces the increase in I_C and improves the operating point stability. Figure 59.4 shows how the rising tendency of I_C is checked by the circuit.

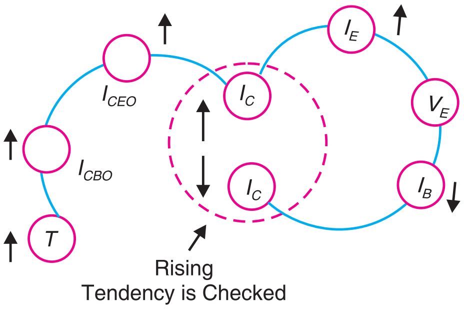


Fig. 59.4

Circuit Analysis. As I_B is very small, it can be taken that same current I_1 is flowing through both the resistors R_1 and R_2. Let us calculate collector current I_C.

$$I_1 = \frac{V_{CC}}{(R_1 + R_2)} \qquad ...(1)$$

The voltage V_2 developed across R_2 is given by

$$V_2 = \left(\frac{V_{CC}}{(R_1 + R_2)}\right) R_2 \qquad ...(2)$$

Applying KVL to the base circuit, we have

$$V_2 = V_{BE} + V_E = V_{BE} + I_E R_E$$

or
$$V_2 = V_{BE} + I_C R_E \qquad (\because I_E \approx I_C)$$

$$\therefore \quad I_C = \frac{V_2 - V_{BE}}{R_E} \qquad ...(3)$$

Here I_C is almost independent of transistor parameters and hence good stabilisation is ensured.

The collector emitter voltage V_{CE} can be calculated as follows :

Applying KVL to the output circuit, we have

$$V_{CC} = I_C R_L + V_{CE} + I_E R_E$$
$$= I_C R_L + V_{CE} + I_C R_E$$
$$\therefore \quad V_{CE} = V_{CC} - I_C (R_L + R_E) \qquad ...(4)$$

Advantages. This biasing method is almost universally used on account of the following advantages :

(*i*) Several combinations of potential divider resistors R_1 and R_2 are possible.

(*ii*) A common power supply V_{CC} is sufficient for providing the bias related to emitter-base and collector-base junctions.

(*iii*) The gain of the transistor is not influenced by negative feedback.

(*iv*) The operating point is stable irrespective of change in temperature.

59.5 Thermal Runaway

The maximum average power $P_{D(max)}$ which a transistor can dissipate depends upon the transistor construction. It may lie in the range from a few *milliwatts* to 200 *W*. This maximum power is limited by the temperature that the collector-to-base junction can withstand. The range of temperature lies between 60°C and 100°C for *Ge* transistor and 150°C and 225°C for *Si* transistor. If the temperature increases beyond this range, then the transistor burns out. The increase in the collector junction temperature is due to *thermal runaway*.

The Effects of Thermal Runaway are Obvious on this Battery Case.

The flow of collector current I_c produces heat within the transistor. This increases the transistor temperature. If the heat is not quickly drawn off, the temperature of the collector-base junction will increase. This increases the minority carrier density. Therefore, the collector leakage current I_{co} increases. This leakage current is highly influenced by temperature changes. It almost doubles for every 6°C rise in temperature in *Ge* and for every 10°C rise in *Si*. Any increase in I_{co} is magnified $(1 + \beta)$ times and increases I_c considerably as seen from equation

$$I_c = \beta I_B + (1 + \beta) I_{co}.$$

This increased collector current further raises the transistor temperature that leads to further increase in I_{co}. If this succession of increases is allowed to continue, collector current may become very large in a matter of seconds, thereby damaging the transistor. **The self-destruction of an unstabilised transistor is known as thermal runaway**.

In order to avoid thermal runaway, some form of stabilisation is necessary. It is important that the operating point is stabilised.

The following two techniques are used for stabilization:

(*i*) *Stabilization techniques*. The technique consists in the use of resistive biasing circuit which permits such a variation of base current I_B so as to maintain I_c almost constant in spite of variation of I_{co}, β and V_{BE}.

(*ii*) *Compensation techniques*. In this technique, temperature sensitive devices such as diodes, transistors, thermistors etc., are used. Such devices produce compensating voltages and currents in such a way that the operating point is maintained stable.

In power transistors, provisions are made to remove heat developed at the junction during working. This is achieved by using *heat sink*, which is a metal sheet connected to the collector and whose surface radiates heat quickly.

59.6 Characteristics of a Transistor Connected in Common Base Configuration

Fig. 59.5 shows the circuit arrangement to draw the characteristics of an NPN transistor connected in Common Base Configuration.

V_{EB} can be changed by varying the position of potentiometer R_1. V_{CB} could be varied by adjusting R_2.

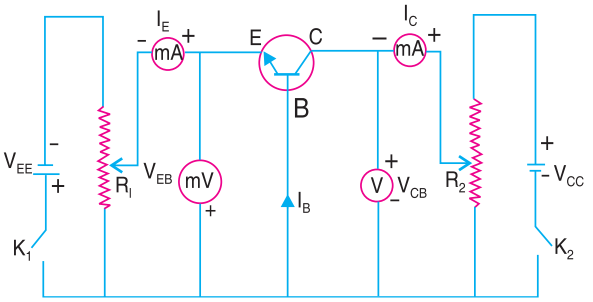


Fig. 59.5

(*i*) Input characteristics : The curve between emitter current I_E and emitter base voltage V_{EB} at constant collector base voltage V_{CB} represents the input characteristic. For plotting the input characteristic the collector base voltage V_{CB} is kept fixed. The emitter base voltage V_{EB} is varied with the help of potential divider R_1 and the emitter current is noted for each value of V_{EB}. A graph of I_E against V_{EB} is drawn. The curve is known as input characteristic. The experiment is repeated for other fixed values of V_{CB}. The input characteristic is shown in Fig. 59.6.

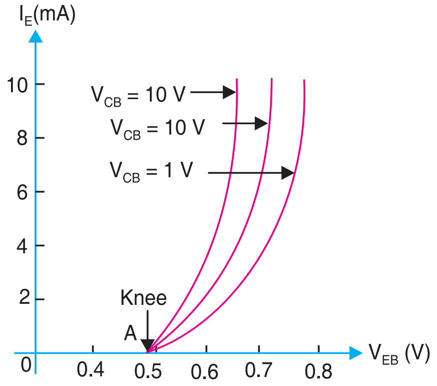


Fig. 59.6

The following points are noted from the characteristic:

(*a*) There exists a threshold voltage V_{EB} (also called the *knee voltage* or *cut-in voltage*), represented by V_K, which is equal to *OA*, below which I_E is negligibly small. $V_K \approx$ 0.5 *V* for *Si* and 0.1 *V* for *Ge* transistors.

(*b*) For $V_{EB} > V_K$, for a constant V_{CB}, I_E increases sharply with a small increase in V_{EB}. This shows that the input resistance is very small.

(*c*) The *dynamic* or *a.c. input resistance* is experimentally obtained from the formula

$$R_i = \left(\frac{\Delta V_{EB}}{\Delta I_E}\right)_{V_{CB}=const.}$$

R_i may be typically of the order of 100 Ω in the linear region.

(*ii*) Output characteristics. The curve between collector current I_C and collector-base voltage V_{CB} at constant emitter current I_E represents the output characteristic.

The emitter current I_E is kept fixed. With the help of potential divider R_2, the value of V_{CB} is varied in steps and the collector current I_C is noted for each value of V_{CB}. Now a graph is drawn between I_C and V_{CB}. The curve so obtained is known as output characteristic. The experiment is repeated for different fixed values of emitter current I_E. Fig. 59.7 shows the output characteristic.

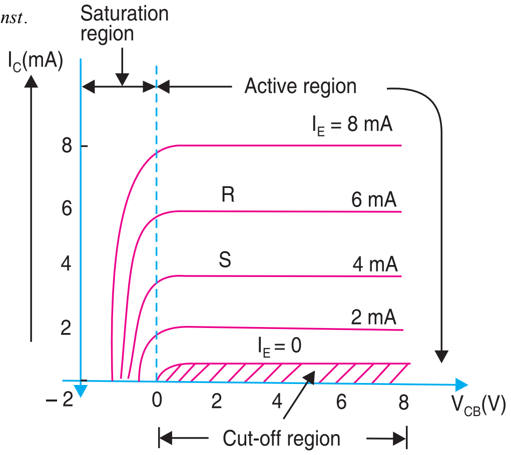


Fig. 59.7

The output set of characteristics has three regions of interest : the *active*, *cutoff*, and *saturation* regions.

(*i*) Active region : It is the region normally employed for linear (*undistorted*) amplifiers.

In the *active region, the base-emitter junction is forward-biased, while the collector-base junction is reverse-biased.*

1. In the active region, I_c = constant ≈ I_E.
2. I_C is practically independent of V_{CB} in the active region. However, if V_{CB} is increased beyond a certain large value, I_C increases rapidly due to *avalanche breakdown* (not shown in the figure) and the transistor action stops.

(*ii*) Cut-off region : In the *cut-off region* (when emitter and collector junctions are both reverse-biased), a small amount of collector current flows even when emitter current $I_E = 0$. This is the collector leakage current I_{CBO}.

(*iii*) Saturation region. The region to the left of ordinate $V_{CB} = 0$ is called saturation region.

1. Note the exponential increase in I_c as the voltage V_{CB} increases towards zero volt.
2. Here, both the $B - E$ and $C - B$ junctions are forward-biased.

Dynamic output resistance (R_0) : It is also called the a.c. *output resistance*.

It is experimentally found from the definition

$$R_0 = \left(\frac{\Delta V_{CB}}{\Delta I_C}\right)_{I_E = \text{const.}}$$

Since the characteristic curves are almost horizontal, R_0 is very high. Its typical value is 500 $k\Omega$.

Current gain. *DC* current gain and *a.c.* current gain are respectively defined by

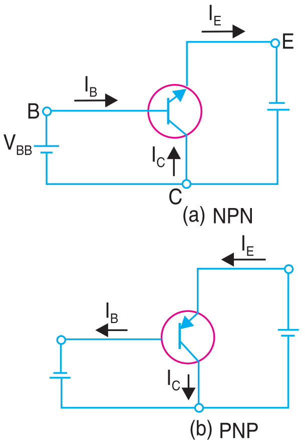
$$\alpha_{dc} = \left(\frac{I_C}{I_E}\right)_{V_{CB}=\text{const.}} \quad \text{and} \quad \alpha_{ac} = \left(\frac{\Delta I_C}{\Delta I_E}\right)_{V_{CB}=\text{const.}}$$

Both α's have a value slightly less than unity.

59.7 C – C Configuration

The *common-collector configuration* is shown in Fig. 59.8 with the proper current directions and voltage notation.

The common-collector configuration is used primarily for impedance-matching purposes since it has a high input impedance and low output impedance.



(a) NPN

(b) PNP

Fig. 59.8

Common Collector Characteristics

Fig. 59.9 shows the circuit arrangement to draw the characteristics of a PNP transistor connected in C – C configuration.

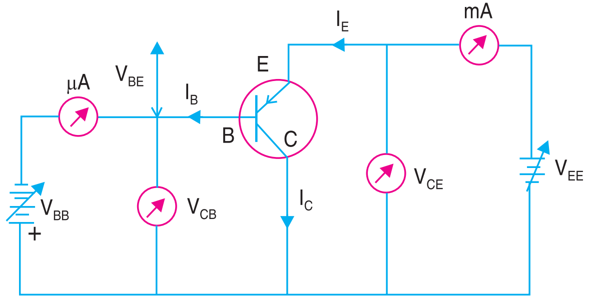


Fig. 59.9

1. Input characteristics : *The curve showing the variation of base current I_B with collector-base voltage V_{CB} at constant collector-emitter voltage V_{CE} is called the input characteristic.*

Fig, 59.10 gives a plot of I_B versus V_{CB} for different values of V_{CE}. These characteristics are quite different from those for *CB* or *CE* circuit because the input voltage V_{CB} is largely determined by the value of *CE* voltage. It may be seen that with an increase in V_{CB}, I_B gets decreased.

2. Output Characteristics : *The curves showing the variation of the emitter current I_E with collector-emitter voltage V_{CE} at constant base current I_B are called the output characteristics of the transistor.*

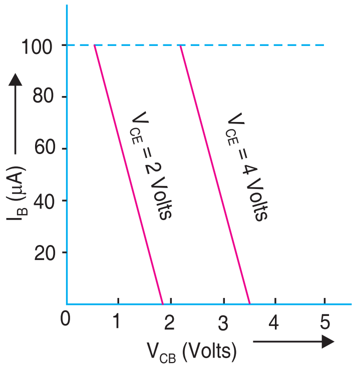


Fig. 59.10

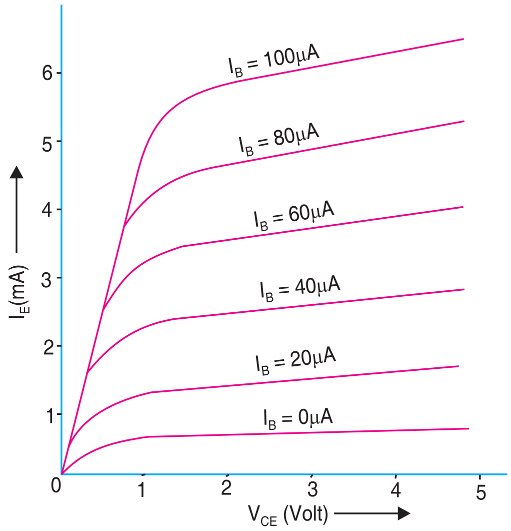


Fig. 59.11

Fig. 59.11 gives the plot of I_E versus V_{CE} for several fixed values of I_B.

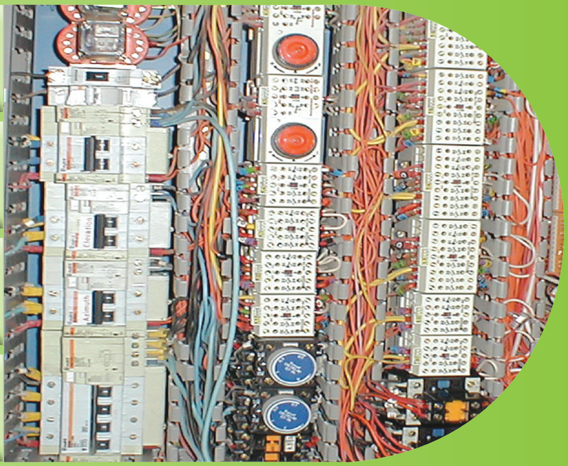
For all practical purposes, the output characteristics of the common-collector configuration are the same as for the common-emitter configuration because $I_C \approx I_E$.

EXERCISE

1. Explain the need for biasing a transistor. What are the different methods of biasing? (*K.U., April* 1991)
2. Why is voltage divider bias commonly used in transistor? Discuss its advantages.

CHAPTER

60



ELECTRONICS II

AT A GLANCE			
60.1	The DC Load Line	**60.2**	A. C. Load Line
60.3	Satellite Communication	**60.4**	Television Camera - The Image Orthicon
60.5	Scanning Process in TV	**60.6**	Colour Television
60.7	Colour Television Camera	**60.8**	Transmission of Colour TV Signal
60.9	Reception of Colour TV Signal	**60.10**	Synchronization in TV
60.11	Television Channel Width	**60.12**	SCR Applications
60.13	Silicon-Controlled Rectifier		

60.1 The D.C. Load Line

When no signal is applied to a transistor circuit it is said to be in the *quiescent condition*. Fig. 60.1 shows a CE amplifier in the quiescent condition.

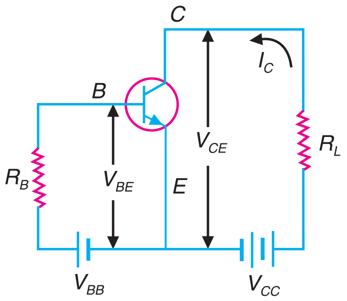


Fig. 60.1

The collector circuit d.c. supply battery V_{CC} sends the collector current I_C through the load resistance R_L and the transistor. The voltage drop across the load is $I_C R_L$. Then the value of collector-emitter voltage V_{CE} at any instant is given by

$$V_{CE} = V_{CC} - I_C R_L \quad \text{...(1)}$$

or
$$I_C = \left(-\frac{1}{R_L}\right) V_{CE} + \frac{V_{CC}}{R_L} \quad \text{...(2)}$$

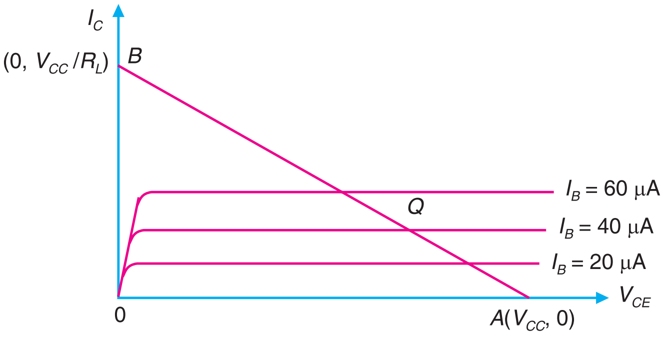


Fig. 60.2

Here $(-1/R_L)$ and (V_{CC}/R_L) are constants. Therefore, if I_C is plotted against V_{CE} (on the transistor's output characteristics), we get a straight line. This line is called the *dc load line* because it depends upon the value of load R_L used in the circuit.

The slope of the line is $(-1/R_L)$. The intercept on the *Y*-axis is (V_{CC}/R_L).

Plotting of DC load line. To draw load line, we need two points of the straight line. They are located as follows :

When collector current $I_C = 0$, then collector-emitter voltage V_{CE} is given by Eq. (1).

$$V_{CE} = V_{CC}. \quad \text{...(3)}$$

Thus, we get point *A* at which $I_C = 0$ and $V_{CE} = V_{CC}$.

When collector-emitter voltage $V_{CE} = 0$, then collector current is given by Eq. (2).

$$I_C = V_{CC}/R_L. \quad \text{...(4)}$$

Thus, we get point *B* at which $I_C = V_{CC}/R_L$ and $V_{CE} = 0$.

Joining points *A* and *B*, we get the d.c. load line (Fig. 60.2).

Location of the Operating Point

The operating point must naturally lie on the load line. To find exactly where on this line it lies we have to consider the base current I_B. The base current is decided by the values of R_B, V_{BB} and V_{BE}. Application of Kirchhoff's voltage law to the base circuit leads us to the equation

$$V_{BB} = I_B R_B + V_{BE} \quad \text{...(5)}$$

The base current is, therefore, given by

$$I_B = \frac{V_{BB} - V_{BE}}{R_B}.$$

Once the base current is known the point where the characteristic for this base current intersects the load line gives the operating point. This point is called the *quiescent operating* point or simply the *Q* point.

60.2 A. C. Load Line

AC Load line is the locus of *Q* point (operating point) when a.c. signal is applied at the input of the amplifier. It enables us to determine the values of collector current and collector-emitter voltage when signal is applied at the input.

The d.c. load line is *AB*. The a.c. load line is *CD* (Fig. 60.3). The intersection of a.c. load line and d.c. load line represents the *Q*-point.

In a.c. load line, a.c. load resistance (R_{ac}) is taken into account, while in d.c. load line, d.c. load resistance (R_L) is considered. DC load line can be drawn as described earlier.

In order to draw a.c. load line, two points on a plot of the output characteristics are required.

Let $(V_{CE})_0$ and $(I_C)_0$ be the zero signal voltage and zero signal collector current corresponding to the operating point, located on the d.c. line.

Let R_L' be the a.c. load resistance (effective load resistance found from a.c. equivalent circuit).

Let $(I_C)_0 \times R_L' = x$. It has the dimensions of voltage.

Let $\frac{(V_{CE})_0}{R_L} = y$. It has the dimensions of current.

Then, for the a.c. load line, the cut-off point on the V_{CE} axis is

$$V'_{CE} = (V_{CE})_0 + x.$$

The saturation current on the I_C axis is

$$I'_C = (I_C)_0 + y.$$

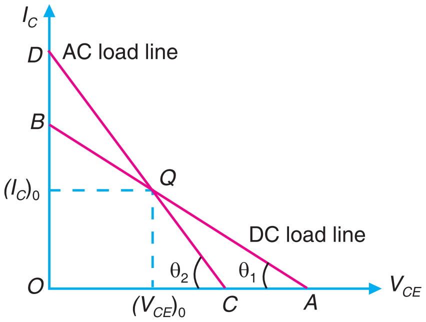


Fig. 60.3

A line drawn with slope $= -1/R_L'$ passing through the Q-point and having $(V'_{CE}, 0)$ and $(0, I'_C)$ as the extreme coordinates gives the a.c. load line for the load resistance R_L'.

Slope of a.c. load line is $\tan\theta_2 = -1/R_L'$.

Slope of d.c. load line is $\tan\theta_1 = -1/R_L$.

60.3 Satellite Communication

Reliable communication links can be made at large distances by using artificial satellites. In this system, an artificial satellite is placed in a circular orbit at an altitude of 36000 km above the earth surface in the equatorial plane. At this orbit, the period of revolution of the satellite around the earth is 24 hrs. Hence, to an observer on the earth, the satellite appears stationary. Such satellites are called *geo-stationary satellites*. The orbit is called the *synchronous orbit*. Three such satellites, 120 degrees apart in the orbit, could then cover the entire earth's surface as shown in Fig. 60.4.

The wave is sent from a transmitter on the earth surface to the satellite. It is processed by the satellite, amplified, and retransmitted towards receiving points on the earth. As the satellite is geostationary, the orientations of the transmitting and the receiving antennas on the earth surface are fixed. The carrier frequency involved is about a few GHz and FM is employed.

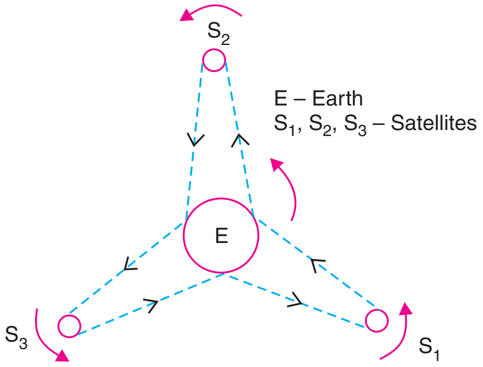


Fig. 60.4

Advantages :

1. Since the carrier frequency for satellite communication is very high (GHz), the antenna size is small and the ionospheric absorption is low.
2. Large number of channels can be formed.
3. The transmitter at the satellite can serve a large area on the earth surface.

Disadvantages :

1. The installation cost is very high.
2. When communication is done through geostationary satellites, due to the large distance involved (approximately 75000 km), there is a large time delay of 250 milliseconds between the transmission and reception of a signal.
3. A satellite once launched and placed in its orbit, the malfunctions in the satellite are highly difficult to correct.

TELEVISION

60.4 Television Camera — The Image Orthicon

A schematic diagram of the image orthicon camera is shown in Fig. 60.5.

It works on the principle of photoelectric emission.

Image orthicon has an image section at its one end and an electron gun and electron multiplier at its other end. Deflecting and focusing coils are fixed at the middle.

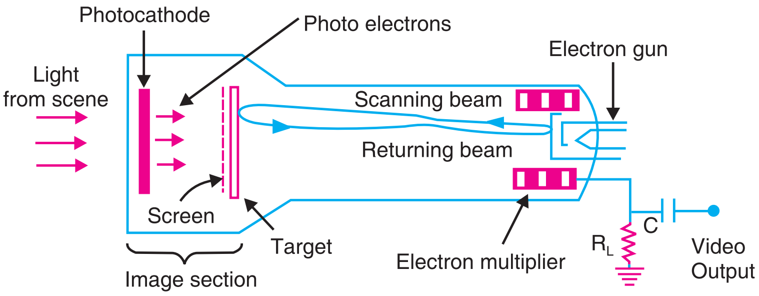


Fig. 60.5

(i) The Image Section. The image section has a photosensitive plate called *photocathode*. It is kept at a negative voltage of the order of – 400 volt. It is so thin that electrons are emitted from its back side in proportion to the light intensity falling on its front side. A *screen* kept at a slightly positive voltage (+ 1 volt), having about 400 meshes per cm and a target mounted close to it are kept at the other end of the image section.

When an optical image of a scene is focussed on the photocathode, photoelectrons are emitted from its back side. These electrons are attracted towards the screen, but most of them pass through the meshes and strike the target. As a result of this, secondary electrons are emitted from the target which are drawn to the screen. Thus, the target is left with a positive charge distribution similar to the optical image focussed on the photocathode.

Television Camera.

(ii) Scanning Section. The back of the target plate is then scanned by an electron beam emitted from the cathode of an electron gun. The number of electrons in the beam is maintained constant. The scanning beam deposits electrons at each point on the back of the target just enough to neutralise the positive charge on the image side. The rest of the electrons in the scanning beam are turned back towards the electron gun. Therefore, the electron density of the returning beam varies in accordance with the positive charge distribution on the target.

***(iii)* Electron Multiplier Section.** The electrons in the returning beam are made to strike the aperture of the electron gun which emits secondary electrons in proportion to the number of electrons in the incident beam. These secondary electrons are then deflected into an electron multiplier system whose output is the amplified reproduction of the returning beam. Thus, a video signal corresponding to the scene projected on the photocathode is available across the load R_L. This video signal is amplified, modulated and transmitted.

60.5 Scanning Process in TV

Scanning is the *process of exploring the scene to be reproduced*. An actual scene, in general, is characterised by a distribution of light and shade over it. The television system at the transmitter breaks a scene into many small elements whose size corresponds to the smallest details of the scene. These are known as the *picture elements*. In the process of scanning, the variations of light intensity are converted into electrical impulses by using the phenomenon of photoelectricity. The electrical impulses of different picture elements are picked up one after another and then transmitted rapidly in sequence through a single channel. At the receiving end, they are reassembled in their proper sequence to create a replica of the original image. Due to rapid scanning and persistence of vision, all the picture elements appear simultaneously before our eyes.

The *resolution or definition* of television is determined by the number of elements into which a scene is broken. Larger the number of elements, better is the definition and finer is the image.

Now a days, the method of *interlaced scanning* is used widely.

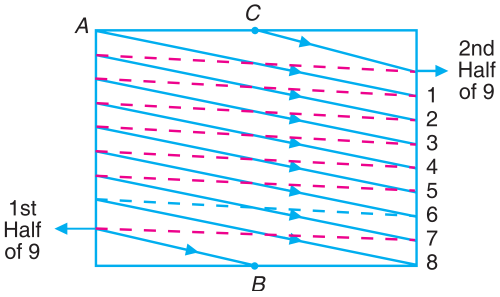


Fig. 60.6

Interlaced Scanning. Schematic illustration of linear interlaced scanning is given in Fig. 60.6. Each picture (frame) is divided into two parts called "fields" and 50 views of the scene are presented to the eye per second. One field consists of odd lines and the other field consists of even lines. Fig. 60.6 shows the interlaced technique of 9-line system. All the odd lines (1st field) in the frame are scanned first from top to bottom skipping over the even lines. After completing all the odd lines, the electron beam retraces back to the top of the frame to start scanning the even lines. Since 25 frames are scanned per second, and each frame is scanned in two fields, the field scanning frequency is 25 × 2 = 50 Hz. By dividing the picture into two fields, interlaced scanning eliminates *flicker*.

60.6 Colour Television

In colour TV system, additional circuits are used for colour transmission and reception with black and white TV system. All the colours can be reproduced by properly combining the three primary colours : *red, green* and *blue*. For transmission of colour pictures, therefore, the object at the transmitting station is televised with three TV cameras, fitted with optical filters, to allow transmission of these three colours respectively. By combining the colour outputs of these cameras, two basic signals are produced for transmission over the 6 MHz channel. These two signals are respectively known as the *luminance signal* and the *chrominance signal*.

The luminance signal. Luminance signal contains only the brightness changes of the picture, as does the ordinary video signal. A conventional black-and-white receiver will receive only this luminance signal in a colour telecast. Therefore, the reproduction will be in black-and-white.

The chrominance signal. The colour information is contained in the chrominance signal. The chrominance signal is a 3.58 MHz signal modulated by colour information.

A colour TV receiver combines the luminance and the chrominance signals and recovers the red, green, and blue video signals in their proper proportion. These signals are amplified and fed to a tricolour picture tube for the display of the coloured image on the viewing screen.

The colour picture tube. A tricolour picture tube, known as *shadow-mask colour tube* is shown in Fig. 60.7. It has three electron guns (red, green and blue) in order to produce three electron beams for primary colours. The amplified chrominance signal which contains red, green and blue colours is applied to corresponding electron guns. The electron guns are so positioned that their beams cross at some distance from the fluorescent screen. A thin metal sheet or *shadow-mask* containing a large number of very fine holes is placed at the plane where the beams cross over. The viewing screen is a glass plate deposited with closely placed groups of phosphor-dot trios. Each trio consists of three phosphor dots which, when hit by electrons, emit red, green and blue light, respectively. A particular hole in the shadow mask is coupled with a particular trio on the screen. The electrons from the "red" gun can strike only red emitting dots. Similarly, the electrons from the "green" and "blue" guns can hit only green emitting dots and blue emitting dots, respectively. If the number of electrons supplied by each gun is separately controlled, the proportions of red, green, and blue in any trio can be varied. Viewing from a large distance from the screen, the eye cannot resolve the individual dots but can only sense a colour given by the proportions of red, green and blue light emitted from the trio. The complete colour picture is produced by scanning the screen with the three beams deflected together.

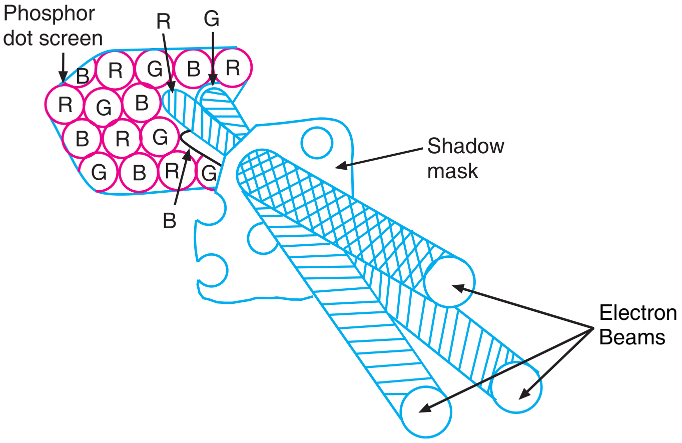


Fig. 60.7

60.7 Colour Television Camera

Fig. 60.8 shows the block diagram of colour TV camera.

Light from the scene, after passing through the lens *L*, is incident on a dichroic mirror system. M_1 and M_2 are total reflecting and M_B and M_R are beam splitting (Partially reflective and partially transmissive).

(*i*) The dichroic mirror M_B reflects blue light and permits red and green lights to pass through it. The blue

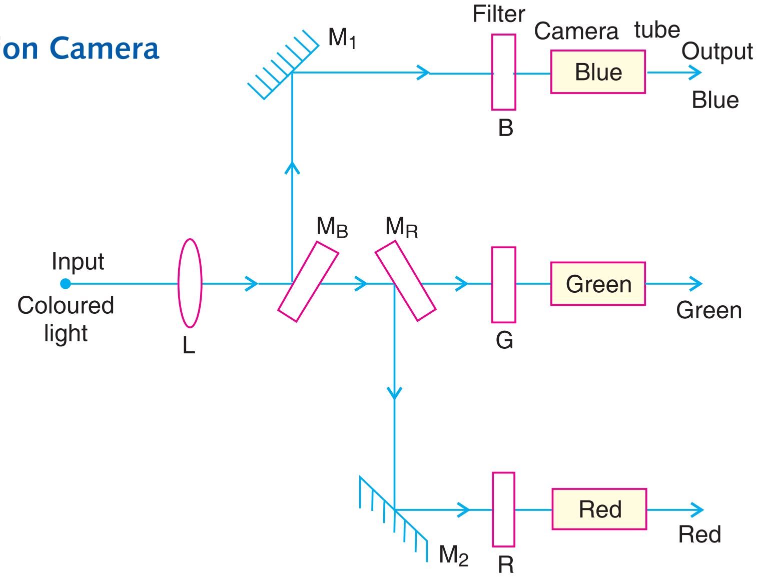


Fig. 60.8

light after reflection at the mirror M_1 enters the blue camera tube via the blue filter. The blue filter allows only blue light, filtering other colours, if any.

(*ii*) The red and green lights from M_B fall on dichroic mirror M_R, which reflects red light and transmits green colour through it. The green light after passing through the green filter, enters the green camera tube.

(*iii*) The red light from M_R gets reflected at the mirror M_2 and passes into the red camera via red filter.

In this way the red, green and blue colour content of a picture is separated into three different light signals.

60.8 Transmission of Colour TV Signal

Fig. 60.9 shows the block diagram of colour TV transmitter.

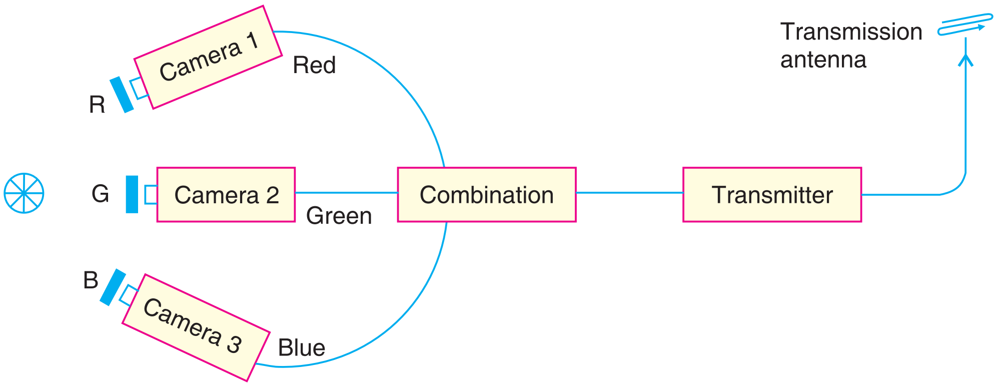

Camera 1
Red
R
Transmission
antenna
G
Camera 2
Green
Combination
Transmitter
B
Camera 3
Blue

Fig. 60.9

The camera system consists of three ordinary television cameras, all focussed on the object to be televised. The red, green and blue filters are placed in front of the cameras 1, 2 and 3. The signals from the three cameras are combined in a special way and telecast together.

60.9 Reception of colour TV Signal

Fig 60.10 shows the block diagram of colour TV Receiver.

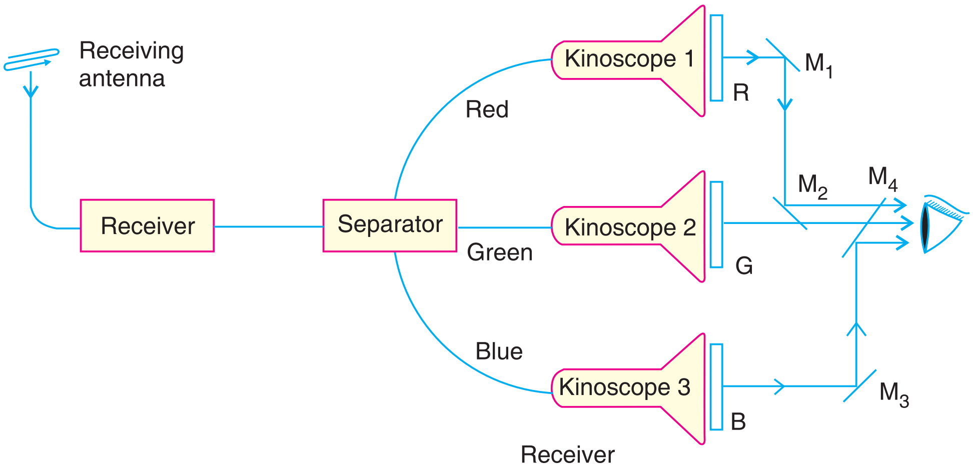

Receiving
antenna
Kinoscope 1
R
M_1
Red
M_2
M_4
Receiver
Separator
Kinoscope 2
Green
G
Blue
Kinoscope 3
B
M_3
Receiver

Fig. 60.10

At the receier, the three signals are received together and then separated. Each signal is used to operate a separate kinoscope in front of whose screen is a second filter of the same colour. The camera behind each colour filter responds to parts of the object having the same colour and therefore the corresponding kinoscope of the receiver shows only that colour. A system of mirrors is used to combine the three coloured beams into a single beam, so that the final image is also coloured like the original object.

60.10 Synchronization in TV

The scanning operations at the transmitter and the receiver must be synchronised or kept in step to reproduce at the receiver the image of the object being televised. This is accomplished by synchronizing signals which are produced at the transmitting station and radiated along with the picture information. These synchronising signals have the form of rectangular pulses and are known as *sync* pulses. They are produced periodically at the line and field scanning frequencies.

For synchronization of horizontal line scanning, horizontal sync pulse is transmitted for each horizontal line. There are 625 horizontal lines in each frame (in Indian system). Each frame is scanned in 1/25 sec.

Total number of horizontal lines scanned per second is 625 × 25 = 15625.

Thus the horizontal synchronizing frequency , $f_h = 15625$ Hz.

Similarly a vertical synchronizing pulse is transmitted for each field to keep the vertical scanning in step. The *vertical synchronizing frequency* is

$$f_v = 2 \times \text{picture repetition rate} = 2 \times 25 = 50 \text{ Hz.}$$

Thus the vertical synchronizing frequency is equal to the field scanning frequency *i.e.*, 50 Hz.

To avoid interference with the picture on the TV screen, the synchronizing pulses are transmitted during the blanking periods *i.e.*, when no scene is being pictured or scanned. Actually, the scanning spot returns quickly to start the next line at the end of each horizontal line. Similarly, the spot flies back to the top of the frame after reaching its bottom. During these retrace or return intervals, the picture tube is made inoperative by means of *blanking pulses*. These *blanking pulses* are negative pulses of large amplitude, produced at the transmitting station and are added to the video signal. These pulses are imposed on the transmitting signals and modulate the transmitter carrier.

At the receiving station, after the detector stage, the blanking pulses are separated from the composite video signal and used to blank the output of the picture tube.

60.11 Television Channel Width

The band of frequencies required to transmit a TV signal is called a "*TV channel*". Television broadcasting stations in India are assigned channel 7 MHz wide, arranged as shown in Fig. 60.11. This includes both picture and sound. VHF and UHF bands are used for transmission of different TV programmes, each having a channel width of 7 MHz. From the transmitter to the receiver, TV signals travel by means of tropospheric or space waves.

In the entire band width of 7 MHz, the video carrier (picture carrier) is placed at 1.25 MHz . The audio carrier (sound carrier) is placed at 6.75 MHz with a maximum deviation of ± 50 kHz. The

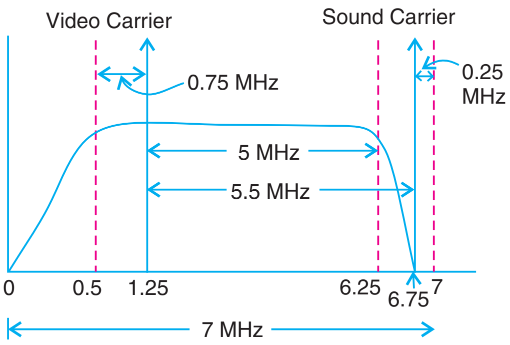


Fig. 60.11

complete lower side band for the picture signal is not transmitted; this is referred to as *vestigial side band transmission*. The width of vestigial side band is 0.75 MHz. The width of the main side band for video is 5 MHz. The upper side band of the amplitude modulated video carrier extends at full amplitude from 1.25 to 6.25 MHz, a range of 5 MHz, and at reduced amplitude upto 6.75 MHz. The sound accompanying the TV picture signal is obtained by a separate FM transmitter having a maximum frequency deviation of 50 kHz and a carrier frequency 5.5 MHz greater than the video carrier. This frequency difference (5.5 MHz) is called the *inter carrier frequency* and is used in extracting the sound information from the video detector.

60.12 SCR Applications

1. Series Static Switch :

Fig. 60.12 shows a half wave *series static switch.*

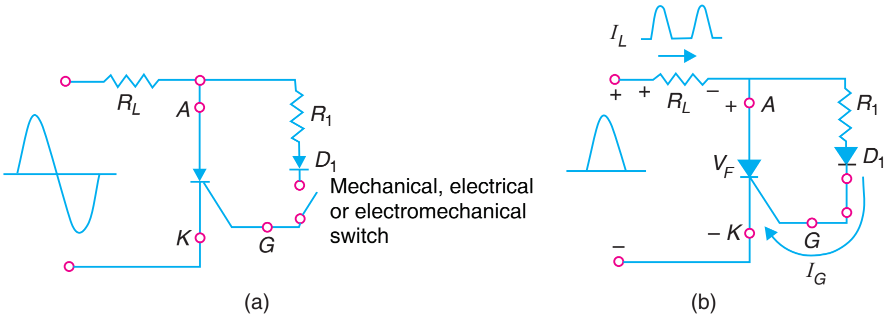


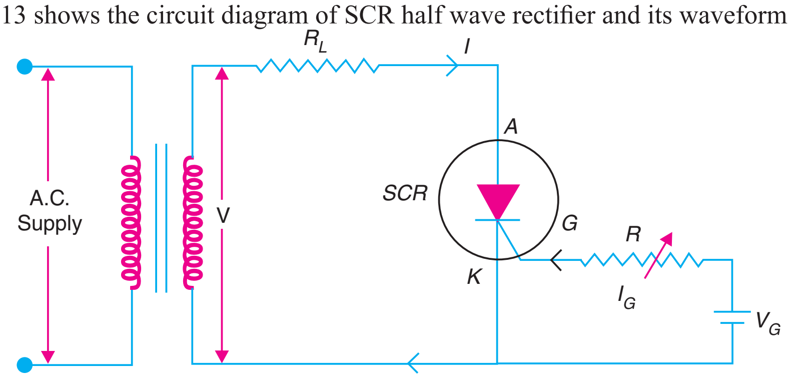
Fig. 60.12

If the switch is closed as shown in Fig. 60.12 (*b*), a gate current will flow during the positive portion of the input signal, turning the SCR on. Resistor R_1 limits the magnitude of the gate current. When the SCR turns on, the anode to cathode voltage (V_F) will drop to the conduction value, resulting in a greatly reduced gate current and very little loss in the gate circuitry. For the negative region of the input signal, the SCR will turn off since the anode is negative with respect to the cathode. The diode D_1 is included to prevent a reversal in gate current.

The waveforms for the resulting load current and voltage are shown in Fig. 60.12 (*b*). The result is a half-wave-rectified signal through the load. If less than 180° conduction is desired, the switch can be closed at any phase displacement during the positive portion of the input signal. The switch can be electronic, electromagnetic, or mechanical, depending on the application.

2. SCR half wave rectifier :

Fig. 60.13 shows the circuit diagram of SCR half wave rectifier and its waveform.



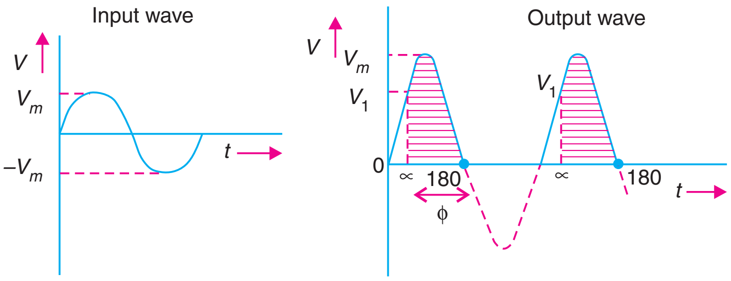


Fig. 60.13

The a.c. supply to be converted into d.c. supply is applied to the primary of the transformer. The load resistance R_L is connected in series with the anode. To control the gate current, a variable resistance R is inserted in the gate circuit. Output is taken across the load R_L.

Operation

(*i*) Suppose the gate current is adjusted to such a value that SCR closes at a positive voltage V_1. So, when the secondary a.c. voltage becomes V_1 in the positive half cycle, SCR conducts. Beyond this voltage SCR continues to conduct till the voltage becomes zero (at which it is turned OFF). SCR conducts current during the conduction angle $\phi = 180 - \alpha$ where α is the firing angle. So we get output voltage not during the entire positive half cycle but from V_1 only.

(*ii*) During the negative half cycle of a.c. voltage appearing across secondary, the SCR will not conduct. Hence output voltage is zero.

Again at the start of the next positive half cycle, when the secondary a.c. voltage becomes V_1, the SCR conducts.

Expression for average current :

Let the alternating voltage appearing across the secondary of the transformer be $V_m \sin\theta$. Suppose the firing angle is α. Then the rectifier will conduct from α to 180° during the positive half cycle.

∴ Average output voltage,

$$V_{av} = \frac{1}{2\pi}\int_{\alpha}^{180^\circ} V_m \sin\theta \, d\theta = \frac{V_m}{2\pi}\int_{\alpha}^{180^\circ} \sin\theta \, d\theta = \frac{V_m}{2\pi}[\cos\alpha - \cos 180^\circ] = \frac{V_m}{2\pi}[1+\cos\alpha]$$

The average current $I_{av} = \dfrac{V_{av}}{R_L} = \dfrac{V_m}{2\pi R_L}(1+\cos\alpha)$

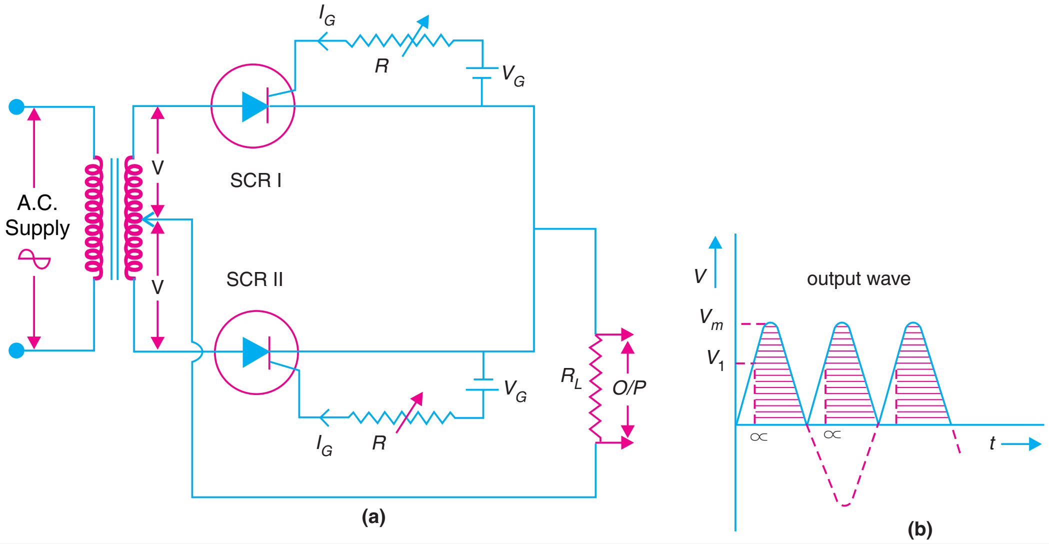
When $\alpha = 0°$, then $I_{av} = \dfrac{V_m}{\pi R_L}$.

When $\alpha = 90°$, then $I_{av} = \dfrac{V_m}{2\pi R_L}$

Thus as the firing angle α increases, I_{av} decreases.

3. SCR full wave rectifier

Fig. 60.14 shows the circuit diagram of SCR full wave rectifier.



(a) (b)

Fig. 60.14

The gates of both SCRs get their supply from two gate controls.

Operation. Suppose the gate currents are so adjusted that SCRs conduct as the secondary voltage (across half winding) becomes V_1. During the + ve half cycle, SCR I conducts. The conduction will start only when the voltage across the upper half of the secondary becomes V_1. During the – ve half cycle SCR II conducts. Current through the load is in the same direction (d.c.) on both half –cycles of input a.c.

Expression for average current in full wave rectifier

Let $V_m \sin\theta$ be the alternating voltage across one end of the secondary and centre tap. Let α be the firing angle.

Average output voltage,

$$V_{av} = \frac{1}{\pi}\int_{\alpha}^{180°} V_m \sin\theta \, d\theta = \frac{V_m}{\pi}[\cos\alpha - \cos 180°] = \frac{V_m}{\pi}(1 + \cos\alpha)$$

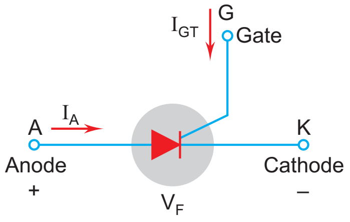
$$\therefore \quad \text{Average current, } I_{av} = \frac{V_{av}}{R_L} = \frac{V_m}{\pi R_L}(1 + \cos\alpha)$$

Thus I_{av} in full wave rectifier is two times that of half wave rectifier.

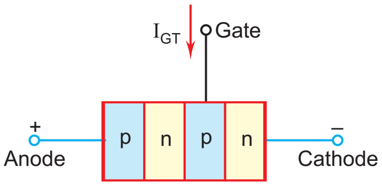
60.13 Silicon-Controlled Rectifier

The SCR is a rectifier constructed of silicon material. Silicon is chosen because of its high temperature and power capabilities.

Construction: It is a four-layered device. It has three terminals. The graphic symbol for the SCR is shown in Fig. 60.15 with the corresponding connections to the four-layer semiconductor structure. As indicated in Fig. 60.15 *a*, if forward conduction is to be established, the anode must be positive with respect to the cathode. This is not, however, a sufficient criterion for turning the device on. A pulse of



(a) SCR Symbol



(b) Basic construction

Fig. 60.15

sufficient magnitude must also be applied to the gate to establish a turn-on gate current, represented symbolically by I_{GT}.

Basic operation of an *SCR*

The four-layer *pnpn* structure of Fig. 60.15 *b* can be split into two three-layer transistor structures [(Fig. 60.16 (*a*)].

Fig. 60.16 (*b*) shows the *SCR two-transistor equivalent circuit.*

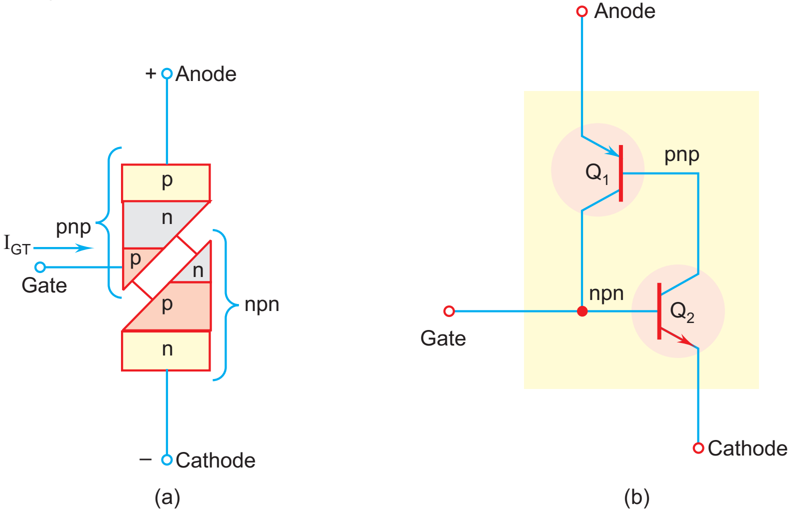
Q_1 is a *pnp* transistor.

Q_2 is an *npn* transistor.

The signal shown in Fig. 60.17 (*a*) is applied to the gate of the circuit of Fig. 60.16 (*b*).

***(i) "Off" state of the SCR.* (When Gate is open)**

During the interval $0 \rightarrow t_1$, $V_{gate} = 0$ V, the circuit of Fig. 60.16 (*b*) will appear as shown in Fig. 60.17 (*b*) ($V_{gate} = 0$ V is equivalent to the gate terminal being grounded).

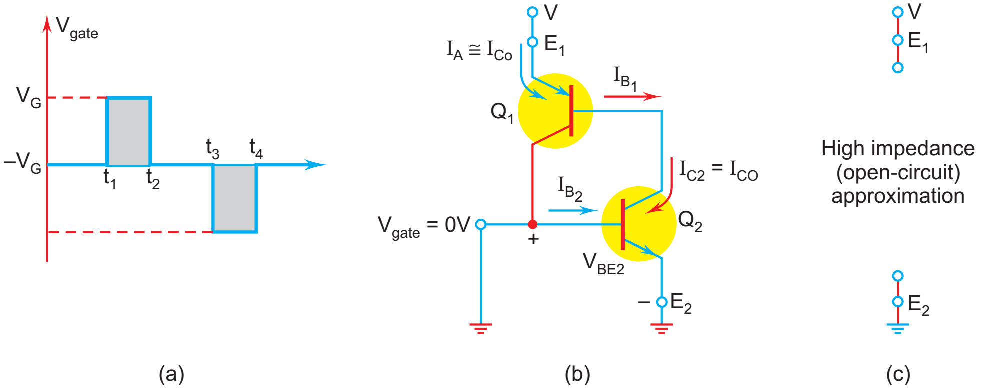

+ Anode
p
n
pnp
I_GT
p
n
Gate
p
npn
n
– Cathode
Anode
pnp
Q1
npn
Q2
Gate
Cathode

(a) (b)

Fig. 60.16

- For $V_{BE_2} = V_{gate} = 0$ V, the base current $I_{B_2} = 0$, and I_{C_2} will be approximately I_{CO}.
- The base current of Q_1, $I_{B_1} = I_{C_2} = I_{CO}$, is too small to turn Q_1 on. Both transistors are therefore in the "off" state, resulting in a high impedance between the collector and the emitter of each transistor.

  Fig. 60.17 (*c*) shows open-circuit representation for the SCR.

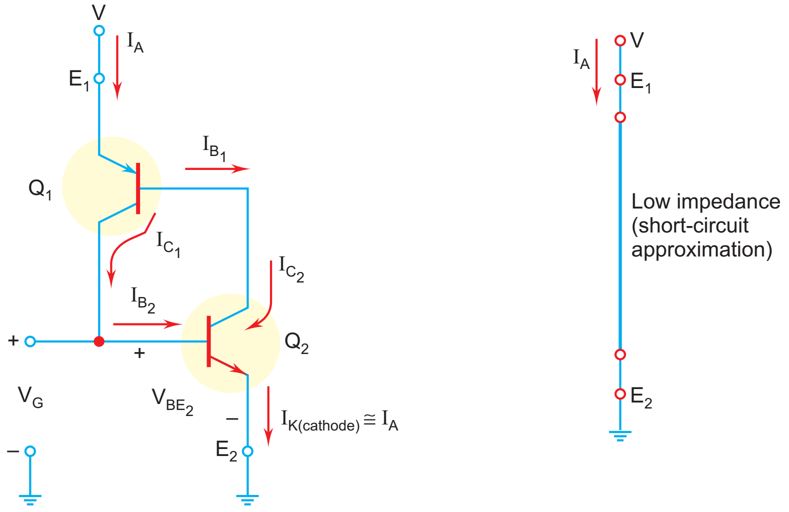

V_gate
V_G
–V_G
t1
t2
t3
t4
V
E1
I_A ≅ I_Co
I_B1
Q1
I_C2 = I_CO
I_B2
Q2
V_gate = 0V
+
V_BE2
– E2
V
E1
High impedance
(open-circuit)
approximation
E2

(a) (b) (c)

Fig. 60.17

(*ii*) *"On" state of the SCR* (When gate is applied +ve voltage w.r.t Cathode)

- At $t = t_1$, a pulse of V_G volts will appear at the SCR gate [(Fig. 60.18 (*a*)]. The potential V_G is chosen sufficiently large to turn Q_2 on ($V_{BE_2} = V_G$). The collector current of Q_2 will then rise to a value sufficiently large to turn Q_1 on ($I_{B_1} = I_{C_2}$).
- As Q_1 turns on, I_{C_1} will increase, resulting in a corresponding increase in I_{B_2}. The increase in base current for Q_2 will result in a further increase in I_{C2}.
- The net result is a regenerative increase in the collector current of each transistor. The resulting anode-to-cathode resistance ($R_{SCR} = V/I_A$) is then small because I_A is large. Fig. 60.18 (*b*) shows the short-circuit representation for the *SCR*.



(a) (b)

Fig. 60.18

- An SCR *cannot* be turned off by simply removing the gate signal. Only a special few can be turned off by applying a negative pulse to the gate terminal at $t = t_3$ [Fig. 60.17 (*a*)].

Volt-ampere (V - I) characteristics of an SCR

The characteristics of an SCR are provided in Fig. 60.19 for various values of gate current. It is the graph drawn between anode - cathode voltage (supply voltage) V_F and the anode current I_A for various values of gate current I_G.

(*i*) Forward characteristics:

- As the supply voltage increases, the anode current remains very small at first and SCR is in the "OFF" state.
- As the supply voltage increases further, the anode current also increases. At a particular voltage ($V_{(BR)F*}$), called *Forward breakover voltage,* SCR is turned "ON" (fired). SCR conducts heavy current.

1. *Forward breakover voltage* $V_{(BR)F*}$ is the voltage above which the SCR enters the conduction region. The asterisk (*) denotes the letter to be added, which is dependent on the condition of the gate terminal as follows:

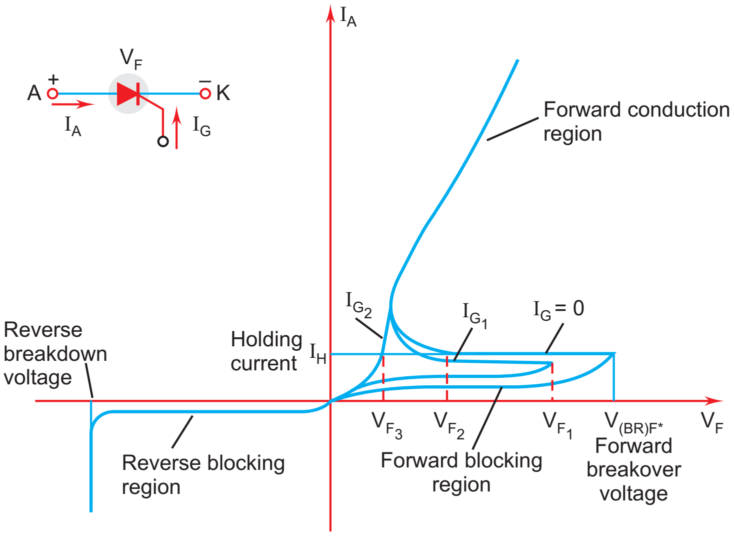


Fig. 60.19

O = open circuit from G to K

S = short circuit from G to K

R = resistor from G to K

V = fixed bias (voltage) from G to K

2. *Holding current* I_H is the value of current below which the *SCR* switches from the conduction state to the forward blocking region under stated conditions.
3. *Forward blocking region* is the region corresponding to the open-circuit condition for the controlled rectifier that *blocks* the flow of charge (current) from anode to cathode.

Curves for varying values of I_G

- For the characteristic having $I_G = 0$, V_F must reach the largest required breakover voltage ($V_{(BR)F*}$) before the "collapsing" effect results and the SCR can enter the conduction region corresponding to the *on* state.
- If the gate current is increased to I_{G1}, by applying a bias voltage to the gate terminal, the value of V_F required for the conduction (V_{F1}) is considerably less. Note also that I_H drops with increase in I_G.
- If the gate current is increased to I_{G2}, the *SCR* will fire at very low values of voltage (V_{F3}). The characteristics will begin to approach those of the basic *p–n* junction diode.

(*ii*) Reverse characteristics:

When anode is given negative voltage w.r.t cathode, the curve between voltage and current is known as reverse characteristic.

- As the reverse voltage increases, the increase in anode current is very small. At a particular reverse voltage, avalanche breakdown occurs and SCR conducts heavily in the reverse direction. This maximum reverse voltage at which SCR starts conducting heavily is known as *"reverse breakover voltage"*.

  Reverse breakdown voltage is equivalent to the Zener or avalanche region of the fundamental two-layer semiconductor diode.
- *Reverse blocking region* is the region correponding to the open-circuit condition for the controlled rectifier that *blocks* the flow of charge (current) from anode to cathode.

CHAPTER

61

HYBRID PARAMETERS

AT A GLANCE

61.1 Two-port Representation of a Transistor
61.2 The Hybrid Equivalent Circuit
61.3 Determination of the h Parameters

61.1 Two-port Representation of a Transistor

A transistor having three terminals is an active device and can be used in any of the three configuration, *CB*, *CE* and *CC*. In these configurations one of the terminals is common to the input and output circuits. Hence a transistor acts as a two port network, input port and the output port (Fig. 61.1).

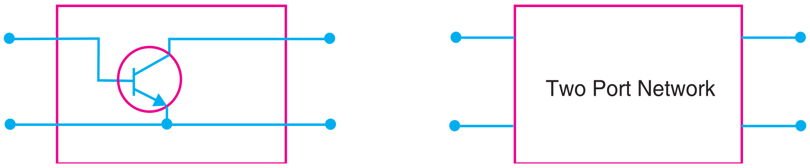


Fig. 61.1

For amplification purposes, the alternating currents, voltages to be amplified (called signals) are applied to the input port (pair of terminals) and the amplified signal is taken from output port (pair of output terminals). Then the performance of the amplifier is obtained by considering the transistor as a 'black box'.

61.2 The Hybrid Equivalent Circuit

To analyse the performance of transistors, they are conveniently represented by an equivalent circuit. The hybrid equivalent or the *h-parameter equivalent circuit* is widely used for small signal low frequency applications because of the following reasons.

(*i*) The *h*-parameters can be measured easily.

(*ii*) They are more independent of each other and other variables like frequency and operating point etc.

(*iii*) The value of *h*-parameters nearly corresponds to actual values of input and output impedances and current gain for many applications.

(*iv*) The *h*-parameters are real numbers at audio frequencies.

(*v*) They are particularly suitable for circuit analysis and design and are specified by the transistor manufacturers.

Hybrid Parameters. A four terminal network or a two port network can be treated as a Black Box with two input terminals and two output terminals (Fig. 61.2).

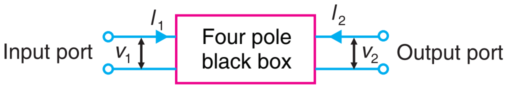


Fig. 61.2

For each pair of terminals, there are two variables, the current *I* and voltage *V*.

These four variables can be related by the following equations :

$$V_1 = h_{11} I_1 + h_{12} V_2 \qquad \text{...(1)}$$

$$I_2 = h_{21} I_1 + h_{22} V_2 \qquad \text{...(2)}$$

The parameters h_{11}, h_{12}, h_{21} and h_{22}, which relate the four variables of the two port system, are called *h-parameters*. They may be defined by first putting $V_2 = 0$ (output terminals short circuited) and then $I_1 = 0$ (Input terminals open circuited) in the above equations. Thus

$$h_{11} = \left|\frac{V_1}{I_1}\right|_{V_2=0} = \text{Input impedance (with output shorted)} = h_i$$

$$h_{21} = \left|\frac{I_2}{I_1}\right|_{V_2=0} = \text{Forward current ratio (with output shorted)} = h_f$$

$$h_{12} = \left|\frac{V_1}{V_2}\right|_{I_1=0} = \text{Reverse voltage ratio (with input open)} = h_r$$

$$h_{22} = \left|\frac{I_2}{V_2}\right|_{I_1=0} = \text{Output admittance (with input open)} = h_0$$

h_{11} and h_{22} (*i.e.*, h_i and h_0) have dimensions of impedance and admittance respectively. h_{12} and h_{21} (*i.e.*, h_r and h_f) have no units. Hence the name hybrid parameters. Another subscript (*b* for Base, *e* for Emitter and *c* for Collector) is added to designate the configuration of the transistor.

Thus, h_{ib}, h_{rb}, h_{fb}, and h_{0b} describe the *CB* configuration, h_{ie}, h_{re}, h_{fe}, h_{0e} and h_{ic}, h_{rc}, h_{fc} and h_{oc} describe respectively the *CE* and *CC* configurations.

Typical values of *h*-parameters for a transistor for different configurations are listed in the following table.

Parameter	*CB*	*CE*	*CC*
h_i (Ω)	30 – 80	600 – 2000	600 – 2000
h_r (ratio)	$5 - 200 \times 10^{-6}$	3×10^{-4}	1
h_f (ratio)	– 0.98	10 – 100	– 51 – 100
h_0 (mhos)	1.5 μ mhos	20 μ mhos	20 μ mhos

61.3 Determination of the h Parameters

We have,
$$V_1 = h_{11} I_1 + h_{12} V_2 \qquad ...(1)$$
$$I_2 = h_{21} I_1 + h_{22} V_2 \qquad ...(2)$$

Determination of h_{ie}

It can be obtained from Eq. (1) by making $V_2 = 0$.

$$V_1 = h_{11} I_1$$

$$h_{11} = \frac{V_1}{I_1} = \text{Input resistance.}$$

h_{11} gives the input resistance of transistor when $V_2 = 0$. V_2 will be zero when we short circuit the output terminals. h_{11} is also called h_{ie} for common emitter mode.

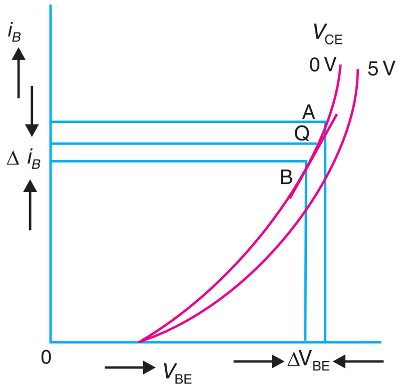


Fig. 61.3

Consider the input characteristics shown in Fig. 61.3. On the characteristics through the operating point Q choose two point A and B on either side of Q and very close to it. On the two axes read the intervals Δi_B and ΔV_{BE} corresponding to AB. Then we have

$$h_{ie} = \frac{\Delta V_{BE}}{\Delta i_B}$$

as V_{CE} is the same at A and B.

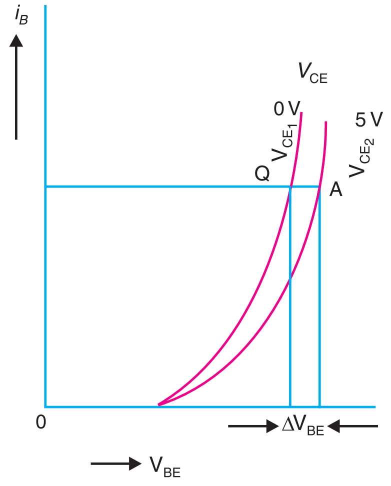


Fig. 61.4

Determination of h_{re}

Putting $I_1 = 0$ in Eq. (1), we get

$$h_{12} = \frac{V_1}{V_2}$$

h_{12} is also called h_{re} for CE mode.

Through the operating point Q draw a line parallel to the V_{BE} axis to cut an adjacent characteristic at A (Fig. 61.4).

$$h_{ie} = \frac{\Delta V_{BE}}{V_{CE_2} - V_{CE_1}}$$

Determination of h_{fe}

Putting $V_2 = 0$ in Eq. (2), we get

$$h_{21} = \frac{I_2}{I_1}.$$

h_{21} is also called h_{fe} for CE mode.

Consider the output characteristics of a transistor (Fig. 61.5).

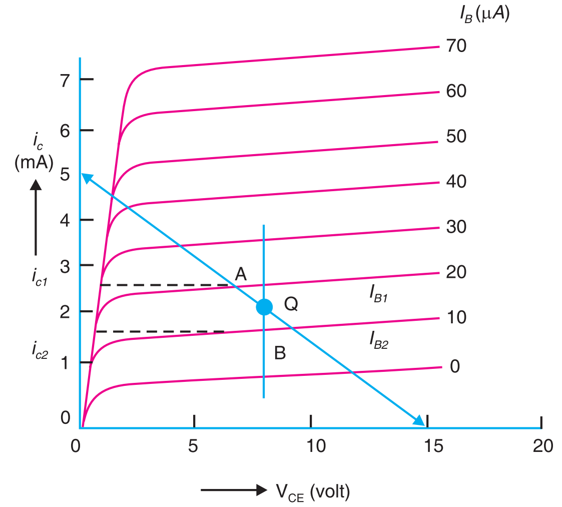


Fig. 61.5

Through the operating point Q draw a line parallel to the i_c axis. Let this line cut the two adjacent characteristics on either side at A and B. Through A and B draw lines parallel to the V_{CE} axis to cut the i_c axis at i_{c1} and i_{c2}. Let the base current corresponding to A be i_{B1} and that corresponding to B be i_{B2}. Then,

$$h_{fe} = \frac{i_{c1} - i_{c2}}{i_{B1} - i_{B2}}$$

Determination of h_{0e}

Putting $I_1 = 0$ in Eq. (2), we get

$$h_{22} = \frac{I_2}{V_2}.$$

h_{22} is also called h_{0e} for *CE* mode.

In relation to the output characteristic curves h_{22} is defined as

$$h_{22} \approx \frac{\Delta i_c}{\Delta V_c} \; (i_B = \text{constant})$$

Through the operating point *Q* draw a line parallel to the characteristics (Fig. 61.6). On this line choose two close points *A* and *B* one on either side of the operating point. The intervals ΔV_{CE} and Δi_c corresponding to *AB* are read on the two axes. Then

$$h_{0e} = \frac{\Delta i_c}{\Delta V_{CE}}$$

as i_B is the same at *A* and *B*.

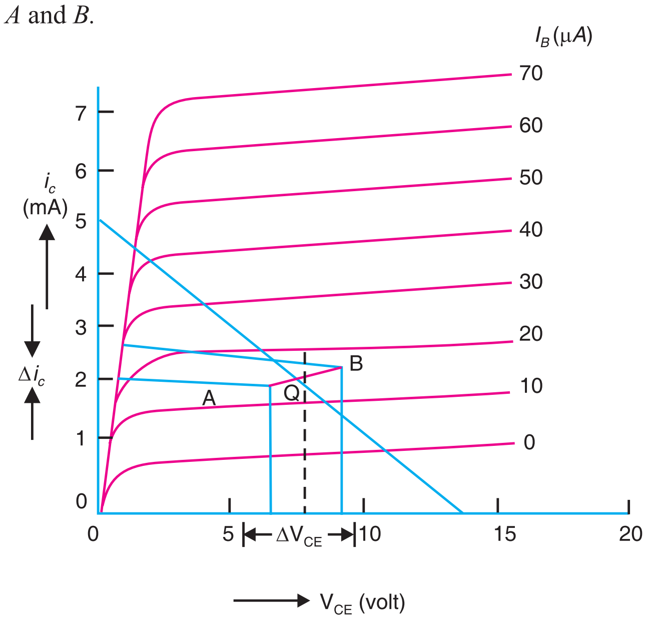


Fig. 61.6

CHAPTER 62

TRANSISTOR AMPLIFIERS

AT A GLANCE

62.1	Hybrid Equivalent Circuit of Common Emitter Amplifier	**62.2**	Analysis of a Transistor CE Amplifier using h-parameters
62.3	Power Amplifier	**62.4**	Class A Power Amplifier
62.5	Transformer Coupled Class A Power Amplifier	**62.6**	Comparison of Amplifier Configurations
62.7	Class B Push-Pull Amplifier		

62.1 Hybrid Equivalent Circuit of Common Emitter Amplifier

Fig. 62.1 shows the common-emitter *NPN* transistor amplifier circuit. R_g is the output resistance of input signal and R_L is the load resistance.

The general *h*-parameter expressions become,

$$V_i = V_{be} = h_{ie} I_b + h_{re} V_o \quad ...(1)$$

$$I_c = h_{fe} I_b + h_{0e} V_0 \quad ...(2)$$

where $V_0 = V_{ce}$.

From Eq. (1), we get

$$I_b = \frac{V_i - h_{re} V_o}{h_{ie}} \quad ...(3)$$

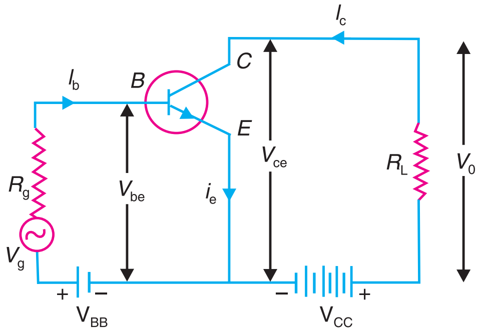


Fig. 62.1

The d.c. voltage of the collector with respect to the emitter is given by

$$V_{ce} = V_{cc} - I_c R_L$$

$$\therefore \quad dV_{ce} = -R_L \, dI_c$$

($\because$ V_{cc} is constant)

or

$$V_{ce} = -R_L \, i_c$$

In terms of usual notations, we can write

$$V_{ce} = -R_L I_c$$

or

$$V_0 = -R_L I_c$$

Substituting the value of V_0 in Eq. (2), we have

$$I_c = h_{fe} I_{\text{b}} - h_{0e} R_L I_c$$

or

$$h_{fe} I_{\text{b}} = h_{0e} R_L I_c + I_c$$

$$= \frac{I_c R_L}{1/h_{0e}} + \frac{I_C R_L}{R_L} \qquad \text{...(4)}$$

Equation (3) indicates that the base-emitter circuit is equivalent to a.c. voltage source of $h_{re} V_0$ which opposes the a.c. input voltage V_i and is connected in series with the input resistance h_{ie}.

Equation (4) indicates that the collector-emitter circuit is equivalent to current source which supplies a current $h_{fe} I_b$ and in parallel of which are connected the output resistance $1/h_{0e}$ and load resistance R_L.

Accordingly, the equivalent circuit is drawn in Fig. 62.2. Here the a.c. voltage source $h_{re} V_0$, which acts in opposition to the input signal V_i, represents the 'feedback' of the output voltage to the input circuit. The current source of magnitude $h_{fe} I_b$ may be looked as if the input current I_b is amplified and appears as $h_{fe} I_b$ in the output circuit. Thus $h_{fe} = \beta$, the current amplification factor.

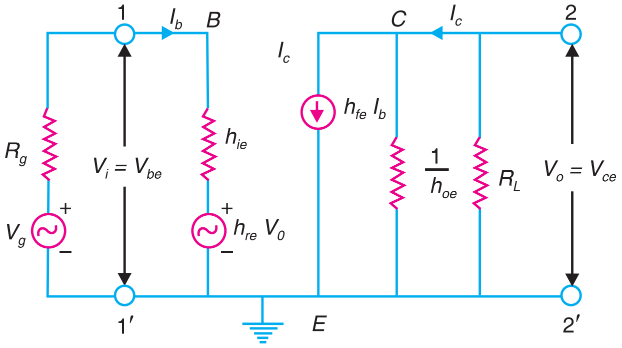


Fig. 62.2

62.2 Analysis of a Transistor CE Amplifier using h-parameters

Figure 62.2 shows the *h*-parameters equivalent circuit of a common emitter transistor amplifier. Here,

h_{ie} = input impedance,
h_{0e} = output admittance,
h_{fe} = forward current gain,
h_{re} = reverse voltage transfer ratio of the transistor.

The signal source V_g is across the input port along with its source impedance R_g. The load resistance R_L appears across the output port. V_i and V_0 are the input and output signals respectively. The input and output currents are taken to be positive, while flowing inward. This circuit is an *a.c.* equivalent circuit and *d.c.* values do not appear in the circuit. I_b and I_c are the input and output currents, with the presence of the source and load.

We will now derive expressions for current gain, voltage gain, input impedance, output impedance and power gain.

(*i*) Current Gain. Let Z be the equivalent impedance of $1/h_{0e}$ and R_L in parallel. Then,

$$\frac{1}{Z} = 1/\frac{1}{h_{0e}} + \frac{1}{R_L} = h_{0e} + \frac{1}{R_L}$$

or
$$Z = \frac{R_L}{1 + h_{0e} R_L}$$

Voltage across R_L = voltage across Z

or
$$I_c R_L = h_{fe} I_b (Z) = h_{fe} I_b \left(\frac{R_L}{1 + h_{0e} R_L} \right)$$

or
$$\frac{I_c}{I_b} = \frac{h_{fe}}{1 + h_{0e} R_L}$$

$$\text{Current Gain } A_{ie} = \frac{\text{Output Current}}{\text{Input Current}}$$

$$\therefore \quad A_{ie} = \frac{I_c}{I_b} = \frac{h_{fe}}{1 + h_{0e} R_L}. \qquad \text{...(1)}$$

(*ii*) Input impedance. The input impedance Z_{ie} of the transistor is the impedance at the input terminals 1 and 1′.

$$\text{Input impedance } Z_{ie} = \frac{\text{Input Voltage}}{\text{Input Current}} = \frac{V_i}{I_b}$$

But
$$\begin{aligned} V_i &= h_{ie} I_b + h_{re} V_0 \\ &= h_{ie} I_b + h_{re} (-I_c R_L) \qquad (\because V_0 = -I_c R_L) \end{aligned}$$

$$Z_{ie} = \frac{V_i}{I_b} = h_{ie} - h_{re} R_L \left(\frac{I_c}{I_b} \right)$$

$$\therefore \quad Z_{ie} = h_{ie} - h_{re} R_L A_{ie} = h_{ie} - \frac{h_{re} \cdot h_{fe} \cdot R_L}{(1 + h_{0e} \cdot R_L)} \qquad \text{...(2)}$$

(*iii*) Voltage gain.

$$\text{Voltage gain } A_{ve} = \frac{\text{Output Voltage} (V_0)}{\text{Input Voltage} (V_i)}$$

But
$$V_0 = -I_c R_L$$

$$A_{ve} = -\frac{I_c R_L}{V_i} = -\left(\frac{I_c}{I_b} \right) \left(\frac{I_b}{V_i} \right) R_L$$

$$= -A_{ie} \left(\frac{1}{Z_c} \right) R_L = -\frac{A_{ie} R_L}{Z_i} \qquad \text{...(3)}$$

Substituting the value of $Z_i = Z_{ie} = h_{ie} - h_{re} R_L A_{ie}$ from Eq. (2),

$$A_{ve} = -\frac{A_{ie} R_L}{h_{ie} - h_{re} R_L A_{ie}} = \frac{R_L}{\frac{h_{ie}}{A_{ie}} - h_{re} R_L}$$

Substituting $A_{ie} = \dfrac{h_{fe}}{1 + h_{0e} R_L}$ from Eq. (1), we get

$$A_{ve} = -\frac{R_L}{\dfrac{h_{ie}(1 + h_{0e} R_L)}{h_{fe}} - h_{re} R_L}$$

$$= -\frac{h_{fe} R_L}{h_{ie}(1 + h_{0e} R_L) - h_{fe} h_{re} R_L}$$

$$= -\frac{h_{fe} R_L}{h_{ie} + (h_{ie} h_{0e} - h_{fe} h_{re}) R_L}$$

$$= -\frac{h_{fe} R_L}{h_{ie} + R_L \Delta h} \qquad ...(4)$$

where $\Delta h = h_{ie} h_{0e} - h_{fe} h_{re}$.

The negative sign shows that the input and the output are 180° out of phase.

(*iv*) Output impedance. The output impedance Z_0 of an amplifier is defined as the ratio of the output voltage to the output current with the input signal generator V_g reduced to zero and replaced by its internal resistance R_g and an a.c. voltage source V_0 (rms) applied to the output terminals as shown in Fig. 62.3. Thus

$$Z_{0e} = \frac{V_0}{I_c}$$

where I_c is the current sent by the applied source.

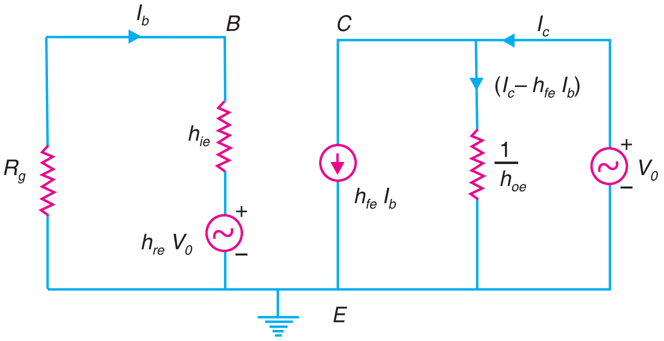


Fig. 62.3

Since the current through the output resistance $1/h_{0e}$ is $I_c - h_{fe} I_b$, the output voltage V_0 is given by

$$V_0 = (I_c - h_{fe} I_b)\frac{1}{h_{0e}}$$

or

$$h_{0e} V_0 = I_c - h_{fe} I_b \qquad ...(5)$$

But the base current I_b is given by

$$I_b = -\frac{h_{re} V_0}{h_{ie} + R_g}$$

Substituting the value of I_b in Eq. (5), we get

$$h_{0e} V_0 = I_c + \frac{h_{fe} h_{re}}{h_{ie} + R_g} V_0$$

or $$V_0\left(h_{0e} - \frac{h_{fe}\,h_{re}}{h_{ie} + R_g}\right) = I_c$$

or $$Z_{0e} = \frac{V_0}{I_c} = \frac{1}{h_{0e} - \dfrac{h_{fe}\,h_{re}}{h_{ie} + R_g}}$$

$\therefore$ $$Z_{0e} = \frac{h_{ie} + R_g}{h_{0e}\,(h_{ie} + R_g) - h_{fe}\,h_{re}} \qquad \text{...(6)}$$

(*v*) Power gain. Power gain of the amplifier is the product of current gain and voltage gain. Thus power gain

$$A_{pe} = |A_{ve}| \times |A_{ie}|$$

Substituting the values of A_{ve} and A_{ie} from Eqs. (4) and (1), we get

$$A_{pe} = \left(\frac{h_{fe}\,R_L}{h_{ie} + R_L\,\Delta h}\right)\left(\frac{h_{fe}}{1 + h_{0e}\,R_L}\right)$$

$$= \frac{h_{fe}^2\,R_L}{(1 + h_{0e}\,R_L)(h_{ie} + R_L\,\Delta h)} \qquad \text{...(7)}$$

where $$\Delta h = h_{ie}\,h_{0e} - h_{fe}\,h_{re}.$$

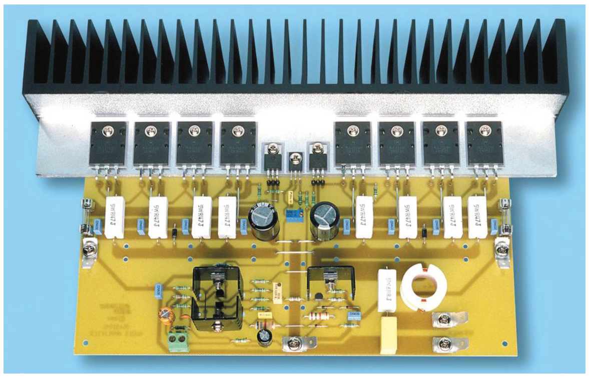
In actual practice, h_{0e}, h_{re} are very small quantities. $h_{0e} < 1$ and $R_L\,\Delta h < h_{ie}$.

$\therefore$ $$A_{pe} = \frac{h_{fe}^2\,R_L}{h_{ie}}.$$

POWER AMPLIFIERS

62.3 Power Amplifier

A transistor amplifier which raises the power level of the signals is known as power amplifier. Here a concentrated effort is made to obtain maximum output power. A power amplifier actually takes power from the *d.c.* supply and delivers it as useful signal power to the load. The ratio of signal power delivered as output to the *d.c.* power input is known as the efficiency of the power amplifier.



Power Amplifier.

In the case of power amplifier, the power drawn by the load impedance is high. This is because of high collector current in this power amplifier. High collector current is the result of low load resistances driven by these amplifiers. These amplifiers (power) are used in the last stage of cascaded amplifier. The last stage drives low resistances like that of a loudspeaker. Power amplifiers use *power transistor* as amplifying device. Power transistor is different from ordinary bipolar transistor in following aspects :

(*i*) Doping level of emitter and base is high.

(*ii*) Ohmic resistance between emitter and base is decreased by increasing the contact area between the base layer and base lead giving it ring like form.

(*iii*) The area of collector is more and collector is often attached to a metallic heat sink in order to dissipate the generated heat.

Power amplifiers are used mainly in four modes — *class A*, *class B*, *class AB* and *class C*.

62.4 Class A Power Amplifier

Figure 62.4 shows a typical *class A transistor power amplifier* operating in the common emitter mode. Here resistors R_1 and R_2 form voltage divider circuit to provide forward biasing of the emitter-base junction. R_E is the emitter resistor for bias stabilisation and C_E is the emitter bypass capacitor which prevents the a.c. voltages from appearing across R_E. The capacitor C_B is called '*Blocking capacitor*'. It is used to block d.c. component of the input signal so that only a.c. signal reaches the base. The amplifier is directly coupled to the load resistance R_L.

Operation of the Amplifier

The collector characteristics of the transistor used in Fig. 62.4 is shown in Fig. 62.5. The load line corresponding to R_L is drawn on the characteristics. The operating point Q is selected in almost the middle of a.c. load line.

When no signal is applied, the Q-point is operating point. When an a.c. signal is applied, the operating point shifts. It keeps on oscillating along the load line with Q as the mean position. The maximum signal condition occurs when the operation swings between saturation and cut-off.

Efficiency. The curves, in Fig. 62.5, show the voltage and current values at saturation and cut-off. In saturation stage, collector current is maximum and collector voltage is minimum. In cut-off stage, voltage is maximum and current is minimum. We denote voltage and current at saturation by $V_{C\,min}$ and $I_{C\,max}$ while these quantities at cut-off are denoted by $V_{C\,max}$ and $I_{C\,min}$ and at Q-point are V_{CQ} and I_{QC}..

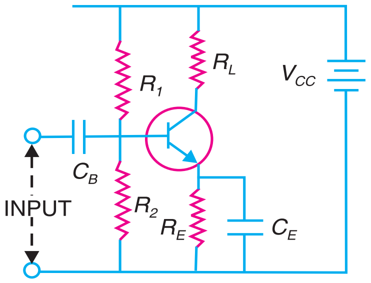


Fig. 62.4

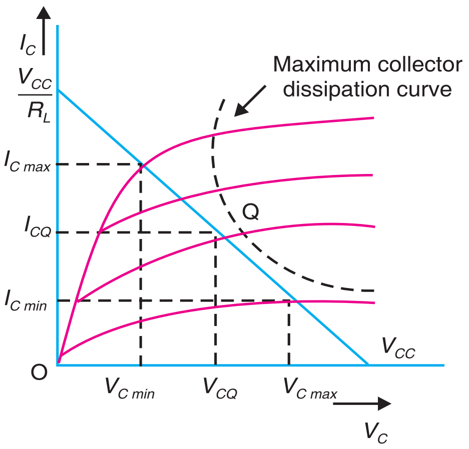


Fig. 62.5

Efficiency η is given by

$$\eta = \frac{\text{output a.c. power}}{\text{input d.c. power}} \qquad \text{...(1)}$$

Input d.c. power is given by

$$P_{in} = V_{cc} I_{CQ} \qquad \text{...(2)}$$

Output a.c. power is given by

$$P_{out} = V_{Crms} I_{Crms} \qquad \text{...(3)}$$

Here V_{Crms} and I_{Crms} are the r.m.s. values of voltage and current. The signal makes excursions between V_{Cmax} and V_{Cmin} (or I_{Cmax} and I_{Cmin}). So, peak values V_{Cm} and I_{Cm} of voltage and current are given by

$$V_{Cm} = \frac{V_{Cmax} - V_{Cmin}}{2}$$

$$I_{Cm} = \frac{I_{Cmax} - I_{Cmin}}{2}$$

Therefore, the a.c. output power

$$\begin{aligned} P_{out} &= V_{Crms} I_{Crms} \\ &= \frac{V_{Cm}}{\sqrt{2}} \times \frac{I_{Cm}}{\sqrt{2}} \\ &= \left(\frac{V_{Cmax} - V_{Cmin}}{2\sqrt{2}}\right) \times \left(\frac{I_{Cmax} - I_{Cmin}}{2\sqrt{2}}\right) \\ &= \frac{(V_{Cmax} - V_{Cmin})(I_{Cmax} - I_{Cmin})}{8} \qquad \text{...(4)} \end{aligned}$$

Therefore, the efficiency of power conversion

$$\eta = \frac{P_{a.c.}}{P_{d.c.}} = \frac{(V_{Cmax} - V_{Cmin})(I_{Cmax} - I_{Cmin})}{8 V_{CC} I_{CQ}} \qquad \text{...(5)}$$

For ideal collector characteristic curves

$$V_{Cmin} = 0, \quad V_{Cmax} = V_{CC}$$

$$I_{Cmin} = 0, \quad I_{Cmax} = 2I_{CQ}$$

$$\therefore \qquad \eta = \frac{(V_{CC})(2I_{CQ})}{8 V_{CC} I_{CQ}} = 0.25.$$

Thus the maximum efficiency obtainable from a class A amplifier when coupled directly to the load resistance is only 25%.

62.5 Transformer Coupled Class A Power Amplifier

Figure 62.6 shows the circuit diagram of a transformer coupled class A power amplifier. R_1 and R_2 form a potential divider arrangement to give base bias. V_{CC} provides reverse bias to the collector-base junction. The transformer is used to couple the transistor to the load. The transformer matches the output impedance of the transistor to the load resistance so that power transfer is maximum.

Transformer Coupled Class A Power Amplifier.

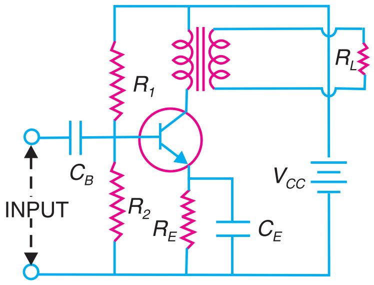


Fig. 62.6

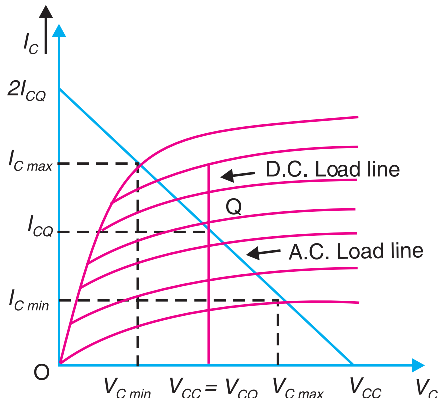


Fig. 62.7

In the circuit shown, load resistance R_L is connected across the secondary coil of the transformer. Its value as seen from the primary coil terminal is given by

$$R_L = n^2 R_L.$$

Here, n is the turns ratio of the transformer given by

$$n = \frac{N_p}{N_s}.$$

Here N_p and N_s are the number of turns in primary and secondary coils, respectively, in practical circuits. Secondary coil also has d.c. resistance, say R_s. The total resistance that as seen at the primary coil terminals is given by

$$R_L' = n^2 (R_s + R_L) \qquad \text{...(1)}$$

Efficiency, $$\eta = \frac{\text{output a.c. power}}{\text{input d.c. power}}.$$

d.c. input power is given by

$$P_{in} = V_{CC} I_{CQ} \qquad \text{...(2)}$$

where I_{CQ} is the collector current at the operating point.

Output a.c. power is given by

$$P_{out} = V_{Crms} I_{Crms}$$

$$= \frac{V_{Cm}}{\sqrt{2}} \cdot \frac{I_{Cm}}{\sqrt{2}}$$

$$= \frac{(V_{Cmax} - V_{Cmin})(I_{Cmax} - I_{Cmin})}{8} \qquad \text{...(3)}$$

From Fig. 62.7, it can be seen that

$$V_{Cmax} = 2V_{CC}, \quad V_{Cmin} = 0$$

$$I_{Cmax} = 2I_{CQ}, \quad I_{Cmin} = 0.$$

$$P_{out} = \frac{4V_{CC} I_{CQ}}{8} = \frac{V_{CC} I_{CQ}}{2}.$$

∴ The efficiency of power amplifier is

$$\eta = \frac{P_{a.c.}}{P_{d.c.}} = \frac{\dfrac{V_{CC} I_{CQ}}{2}}{V_{CC} I_{CQ}}$$

$$\therefore \qquad \eta = 0.5$$

∴ The efficiency of transformer coupled class *A* amplifier is 50% in ideal case.

62.6 Comparison of Amplifier Configurations

The choice of configuration and the type of transistor used depends on the desired input and output impedances, current and voltage gains and the frequency response. The following table gives a comparison of small signal amplifiers in the three possible configurations.

Parameter	*Common base*	*Common emitter*	*Common collector*
Voltage gain	High	Very-high	nearly 1
Current gain	nearly 1	high	highest
Input impedance	lowest	moderate	highest
Output impedance	highest	moderate	lowest
Phase reversal	no	yes	no

62.7 Class B Push-Pull Amplifier

Fig. 62.8 shows the circuit arrangement of a class *B* push-pull amplifier.

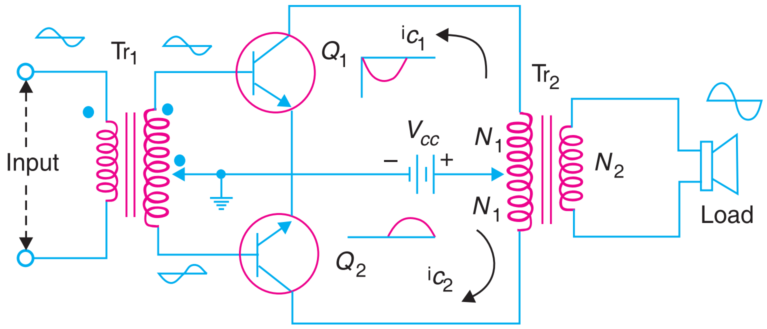


Fig. 62.8

Class B Push-Pull Amplifier.

It consists of two identical NPN transistors Q_1 and Q_2 placed in parallel (back to back) and operated as class *B* amplifiers. The input transformer Tr_1 has a centre tapped secondary winding which provides opposite polarity signals to the two transistors. So the input signal drives the two transistors on alternate half cycles. The collector terminals of the transistors are connected across the primary of the output transformer Tr_2. The *dc* supply voltage V_{cc} is applied at the centre tap of this primary. The load resistance (loudspeaker) is connected across the secondary of the output transformer. When no signal is applied, both the transistors are cut-off.

Working :

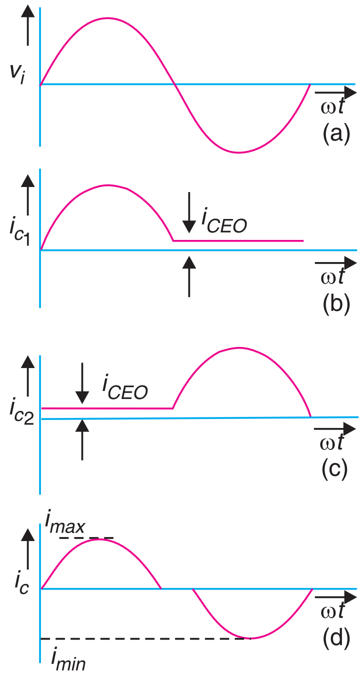


Fig. 62.9

(*i*) When the input signal is applied, the phase splitter transformer Tr_1 produces two signals which are 180° out of phase with each other. The transistors Q_1 and Q_2 are driven by these two signals.

(*ii*) During positive half cycle of the signal, transistor Q_1 conducts because its base is driven positive. Now collector current i_{c_1} flows as shown in Fig. 62.8. Transistor Q_2 does not conduct because its base has negative voltage. Thus i_{c_2} is zero. In this way one positive half cycle of ouput signal appears across the primary of Tr_2.

(*iii*) During the nagative half cycle of the input signal, Q_2 is forward biased and allows a current i_{c_2} to flow. Transistor Q_1 becomes nonconducting.

Thus only one transistor conducts at a time.

The output transformer serves to join the two currents producing an almost undistorted output waveform as shown in Fig. 62.9.

(*i*) Fig. 62.9 (a) shows the input signal v_i.

(*ii*) The collector currents in Q_1 and Q_2 transistors are represented by Figs. (*b*) and (*c*).

(*iii*) The current responsible for output voltage across the load is represented by Fig. (*d*).

Cross-Over Distortion :

When the *d.c.* base voltage is zero, both transistors are off. In order that a transistor begins to conduct, the input signal voltage must have exceeded V_{BE} (Fig. 62.10). The transistor remains cut off until the base voltage exceeds 0.3 *V* for Ge and 0.7 *V* for Si.

Because of this there will be a time interval between the positive and negative alternations of the input, when neither transistor is conducting. The output does not follow the input around the zero input voltage condition. Thus the output collector current is not a uniformly enlarged sine wave at low values of current. The resulting distortion in the output waveform is called crossover distortion.

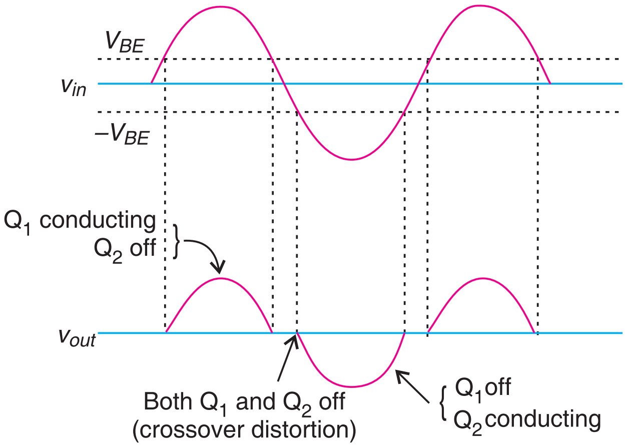


Fig. 62.10

Power Efficiency of Class B Push-Pull Amplifier

The current in each transistor is the average value of half sine loop.

For half sine loop, I_{dc} is given by

$$I_{dc} = \frac{(I_c)_{\max}}{\pi}$$

$$\therefore \quad (P_{in})_{dc} = 2 \times \left[\frac{(I_c)_{\max}}{\pi} \times V_{CC}\right] \qquad \text{...(1)}$$

Here factor 2 is introduced as there are two transistors in push-pull amplifier.

R.M.S. value of collector current = $(I_c)_{\max} / \sqrt{2}$.

R.M.S. value of output voltage = $V_{CC} / \sqrt{2}$.

(Under ideal conditions of maximum power)

$$\therefore \quad (P_0)_{ac} = \frac{(I_c)_{\max}}{\sqrt{2}} \times \frac{V_{CC}}{\sqrt{2}} = \frac{(I_c)_{\max} \times V_{CC}}{2} \qquad \text{...(2)}$$

Now overall maximum efficiency

$$\eta_{\text{overall}} = \frac{(P_0)_{\text{a.c.}}}{(P_{\text{in}})_{\text{d.c.}}}$$

$$= \frac{(I_c)_{\max} \times V_{CC}}{2} \times \frac{\pi}{2(I_c)_{\max} \times V_{CC}}$$

$$= \frac{\pi}{4} = 0.785 = 78.5\%$$

Advantages. (*i*) The output has much less distortion due to cancellation of even harmonics.

(*ii*) The maximum efficiency of class *B* push pull amplifier is quite high (≈ 79%).

EXAMPLE 1. *Prove that a push-pull circuit balances all the even harmonics in the output and leaves the third harmonic term as the principal source of distortion.*

(Rohilkhand, 1999, Kumaun, 2000)

SOLUTION. A power amplifier has to handle large signal inputs. Hence there will be harmonic distortions (frequency components in the signal getting modified) due to non-linear characteristics of the transistors used in the amplifier.

Let the input signal be sinusoidal. The base currents in the two transistors will be opposite in phase. Hence the collector currents will also be opposite in phase.

The collector currents in transistors Q_1 and Q_2 are given by

$$i_{c_1} = A_0 + A_1 \sin \omega t + A_2 \sin 2\omega t + A_3 \sin 3\omega t + \ldots, \qquad \text{... (1)}$$

and

$$i_{c_2} = A_0 + A_1 \sin(\omega t + \pi) + A_2 \sin 2(\omega t + \pi) + A_3 \sin 3(\omega t + \pi)$$

or

$$i_{c_2} = A_0 - A_1 \sin \omega t + A_2 \sin 2\omega t - A_3 \sin 3\omega t + \ldots \qquad \text{... (2)}$$

The total current through the primary of output transformer is

$$i = i_{c_1} - i_{c_2}$$

$$\therefore \quad i = 2[A_1 \sin \omega t + A_3 \sin 3\omega t + \ldots] \qquad \text{... (3)}$$

The voltage induced in the secondary of the output transformer is proportional to *i*.

$$\therefore \quad v_0 \propto i$$

$$v_0 = ki$$ where *k* is the constant of proportionality.

$$\therefore \quad v_0 = 2k[A_1 \sin \omega t + A_3 \sin 3\omega t + \ldots] \qquad \text{... (4)}$$

Thus, the output is free from even harmonics.

The principal source of distortion is the third harmonic 3ω.

CHAPTER

63

SMALL SIGNAL AMPLIFIERS

AT A GLANCE

63.1 Resistance-capacitance (R-C) Coupled Amplifiers
63.2 Transformer Coupled (TC) Amplifier
63.3 Emitter Follower
63.4 Noise in Electronic Circuits
63.5 Distortion in Amplifiers

63.1 Resistance-Capacitance (R-C) Coupled Amplifiers

Fig. 63.1 shows the circuit arrangement of a two stage *RC* coupled *CE* mode transistor amplifier.

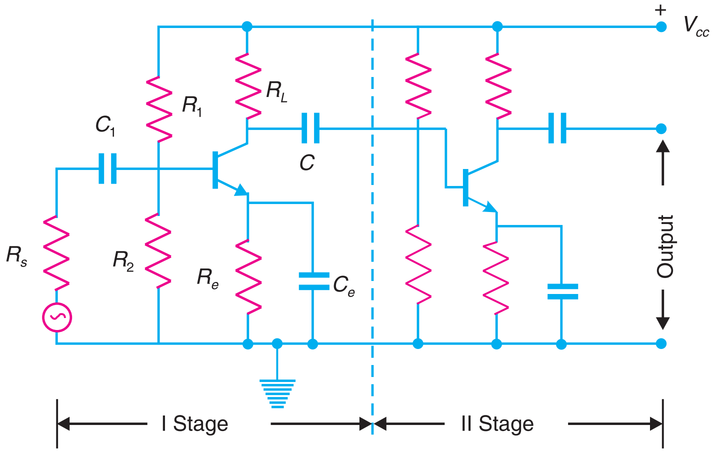


Fig. 63.1

Function of Circuit Components

(*i*) Biasing and stabilization network. The resistances R_1, R_2 and R_e form the biasing and stabilization network. The biasing circuit establishes a proper operating point. Otherwise, a part of the negative half cycle of the signal e_s may be cut off in the output. R_e also reduces the variation of collector current with temperature.

(*ii*) Input capacitor C_1. Electrolytic capacitor C_1 is used to couple the a.c. signal voltage to the base of the transistor. If it is not used, the resistance of the signal source will be in parallel to R_2 and the bias voltage of the base will change.

(*iii*) Emitter bypass capacitor C_e. The capacitor C_e connected across the emitter resistor R_e is of large value. Its reactance $1/\omega C_e$ is small. It, therefore, offers a low reactance path to the alternating component of emitter current and thus bypasses resistor R_e at audio frequencies. Consequently, the potential difference across R_e is due to the d.c. component of the current only. If it is not used, the amplified a.c. signal flowing through R_e will cause a voltage drop across it, thereby reducing the output voltage.

(*iv*) Coupling capacitor C. The coupling capacitor C transmits a.c. signals but blocks d.c. It couples the output of the first stage amplifier to the input of the second stage. It blocks the d.c. voltage of the first stage from reaching the base of the second stage. Thus the d.c. biasing of the next stage is not disturbed. That is why, the capacitor C is also called a *blocking capacitor*.

Operation. When an a.c. signal is applied to the base-emitter circuit of first transistor, it appears in the amplified form across the load resistor R_L. The amplified signal is applied between base and emitter circuit of second stage through coupling capacitor C. The function of coupling capacitor is to allow the passage of amplified a.c. signal and block the passage of d.c. currents and voltages from I stage to the input of the II stage. The second stage does further amplification of the signal. Thus the amplifiers connected in cascade arrangement amplify the signal and the overall gain is considerably increased.

Frequency Response Curve

The voltage gain of the amplifier depends on the frequency of a.c. signal applied. The curve representing the variation of gain of an amplifier with frequency is called frequency response curve. The frequency response curve of RC coupled amplifier is shown in Fig. 63.2.

It is divided into three regions.

The voltage gain is uniform in the midfrequency range.

The voltage gain falls off in both the low frequency and high frequency ranges.

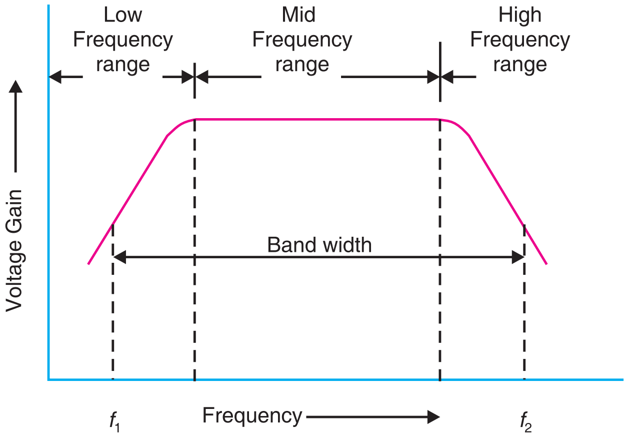


Fig. 63.2

This behaviour is explained below :

(*i*) Low Frequency Range. At low frequencies (< 50 Hz), the *gain decreases with decreasing frequency due to coupling capacitor* C. As the frequency gets lower and lower, the reactance $(1/\omega C)$ of the coupling capacitor gets higher and higher. Consequently, an increasing portion of the voltage drop takes place across C. Thus the voltage reaching from one stage to the next stage decreases. Hence the voltage gain drops.

The *other component due to which the gain decreases at low frequencies is the bypass capacitor* C_e. As frequency decreases, the reactance of capacitor C_e increases and becomes comparable to emitter resistance R_e. Therefore, it cannot shunt the resistor R_e effectively.

Because of these two reasons, the voltage gain decreases in the low frequency region.

Lower cut off frequency f_1. *It is defined as the frequency at which the magnitude of the voltage gain in the low frequency range falls to* $1/\sqrt{2}$ *of the mid-frequency gain.*

(*ii*) Mid Frequency Range. In the middle frequency range (50 Hz to 20 kHz), the *voltage gain of the amplifier is uniform* because resistor values are independent of frequency changes. As the frequency increases in this range, reactance of coupling capacitor decreases which tends to increase the gain. But it simultaneously increases the loading effect of the first stage which decreases the gain. These two factors cancel each other. Hence the gain at mid frequencies remains constant.

(*iii*) High Frequency Range. In the high frequency range (> 20 k Hz), the *voltage gain falls due to the capacitances that appear in shunt across the output*. These capacitances consist of the collector capacitance of the transistor and the stray wiring capacitances. With increase in frequency, the reactance of these capacitances decreases and an increased part of the output is bypassed to earth. Consequently the output voltage of the amplifier decreases. Hence the voltage gain falls at high frequencies.

Moreover, at high frequencies, capacitive reactance of base-emitter junction is reduced. As a result, the base current increases. Hence the current amplification factor β decreases. This has the effect of further reducing the voltage gain.

Upper cut-off frequency f_2. *It is defined as the frequency at which the magnitude of the high frequency gain is* $1/\sqrt{2}$ *times the magnitude of the mid frequency gain.*

Bandwidth. The frequency difference $f_2 - f_1$ gives a measure of the *bandwidth* of the amplifier.

Expression for the voltage gain in mid-frequency range

We make the following simplifying assumptions :

1. The transistors used in the two stages are identical. The input resistance h_{ie} and forward current transfer ratio h_{fe} of each is the same.
2. The output admittance h_{oe} and the reverse voltage gain h_{re} are very small. Therefore they are neglected.
3. The bias resistors R_1, R_2 are of sufficiently large value so that they do not affect the a.c. operation of the circuit.
4. The reactance of the emitter bypass capacitor is negligibly small.

In middle frequency range, the reactance of the coupling capacitor *C* is negligibly small. Fig. 63.3 shows the simplified hybrid equivalent circuit of the first stage of *RC* coupled amplifier.

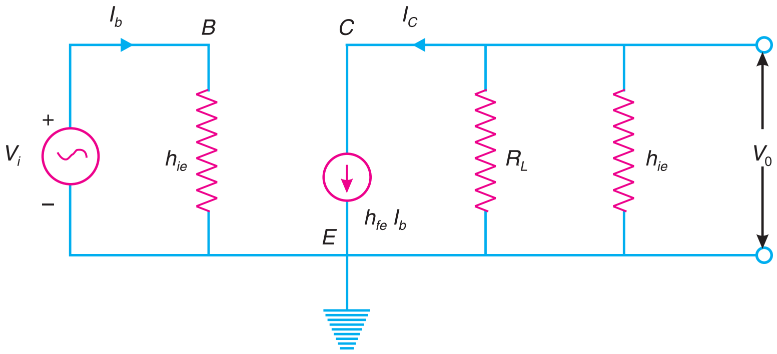


Fig. 63.3

The effective load resistance Z_L is the parallel combination of load resistance R_L and the input resistance h_{ie} of the second stage. Thus

$$\frac{1}{Z_L} = \frac{1}{R_L} + \frac{1}{h_{ie}}$$

or

$$Z_L = \frac{h_{ie} R_L}{h_{ie} + R_L} \qquad ...(1)$$

Therefore, the output voltage V_0 across Z_L is

$$V_0 = -h_{fe} I_b Z_L \qquad ...(2)$$

The input voltage V_i is

$$V_i = h_{ie} I_b \qquad ...(3)$$

Therefore, the voltage gain is given by

$$A_{vm} = \frac{V_0}{V_i} = -\frac{h_{fe} I_b Z_L}{h_{ie} I_b} = -\frac{h_{fe} Z_L}{h_{ie}}\,.$$

Substituting the value of Z_L from Eq. (1), we get,

$$A_{vm} = -\frac{h_{fe}}{h_{ie}}\left(\frac{h_{ie} R_L}{h_{ie} + R_L}\right) = -\frac{h_{fe} R_L}{h_{ie} + R_L} \qquad ...(4)$$

The negative sign in Eq. (4) shows that there is a phase shift of 180° between the output and input signal voltages.

It is evident from Eq. (4) that the voltage gain of *RC* coupled common-emitter amplifier in mid-frequency range is independent of frequency.

63.2 Transformer Coupled (TC) Amplifier

Fig. 63.4 shows the circuit of a two stage transformer coupled amplifier. T_1 is the coupling transformer. It is used to couple the two stages. T_2 is the output transformer. It is used to transfer the output of the second stage to the load. The primary *P* of the coupling transformer T_1 serves as the collector load. Its secondary *S* gives output to the next stage.

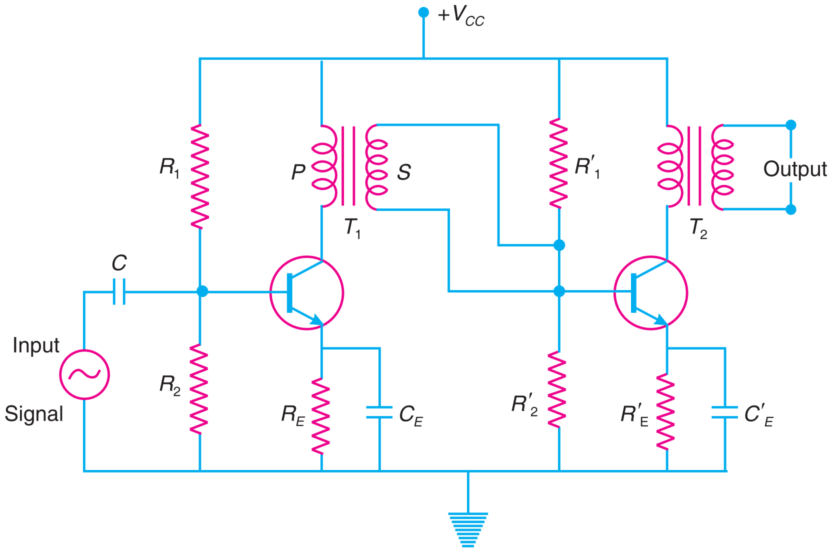


Fig. 63.4

Transformer Coupled Amplifier.

The resistors R_1, R_2 as well as R_1' and R_2' form voltage-divider circuits. The resistors R_E and R'_E are the emitter stabilisers. The capacitors C_E and C'_E are the bypass capacitors. C is the input coupling capacitor.

Operation

The signal is applied to the base of the first transistor. It appears in an amplified form across the primary P of the coupling transformer T_1. This amplified voltage is transferred to the secondary of the transformer by magnetic induction. From there, it is applied to the base of second transistor. Thus the amplified output of first stage becomes the input of the second stage.

The second transistor further amplifies the input. The amplified signal appears across the primary of the output transformer T_2. The secondary of this transformer transfers this output to the load.

Frequency Response

Fig. 63.5 shows the frequency response curve of a transformer coupled amplifier.

(*i*) The frequency response is poor *i.e.*, the gain is constant only over a small range of frequency.

(*ii*) There is a decrease in gain at low frequencies.

(*iii*) At high frequencies, the distributed capacitance between different turns of windings acts as a bypass capacitor. The output voltage decreases. Hence the gain decreases.

There is a peak or exaggerated gain at resonant frequency f_0 of the tuned circuit which is formed by inductance and distributed winding capacitance in the circuit.

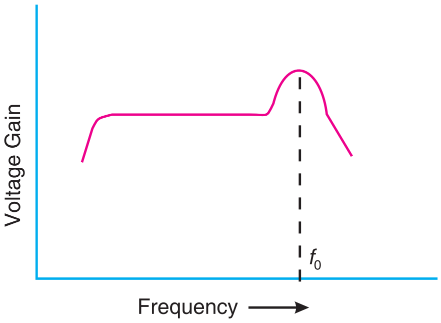


Fig. 63.5

Thus all frequencies in a composite signal, such as music, speech etc., are not equally amplified. Hence transformer coupled amplifier shows *frequency distortion*.

Advantages

1. The operation of a transformer coupled amplifier is more efficient than an RC coupled system because of the low d.c. resistance of the primary which is connected to the collector circuit.

2. By suitably adjusting the turns ratio of the transformer, very good impedance matching can be obtained. Hence very good overall gain can be obtained.

3. No signal power is lost in the collector or base resistors.

Disadvantages

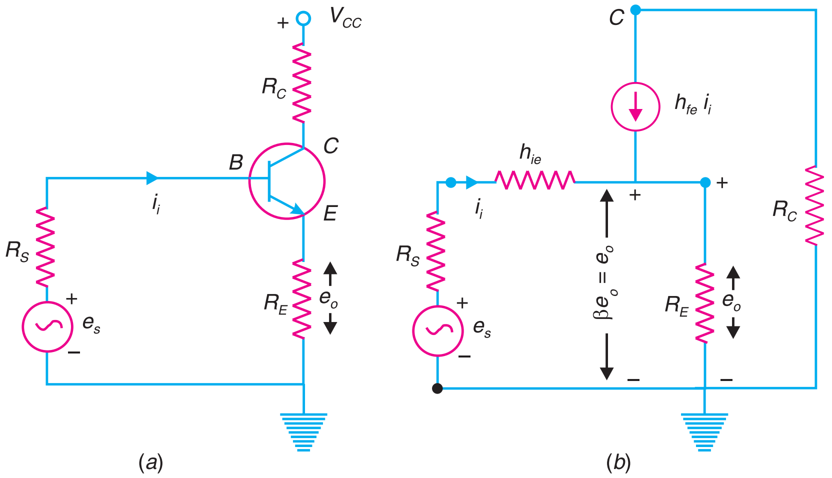
1. The frequency response is poor.
2. The coupling and output transformers are bulky and costly.
3. At radio frequencies, the inductance and interwinding capacitance in the transformer create problems.
4. The transformer windings tend to produce a *hum* in the output.

Applications

A transformer coupled amplifier is mostly employed for impedance matching and power amplification.

63.3 Emitter Follower

Fig. 63.6 (*a*) shows the emitter follower circuit. It is an excellent example of an amplifier with negative feedback.

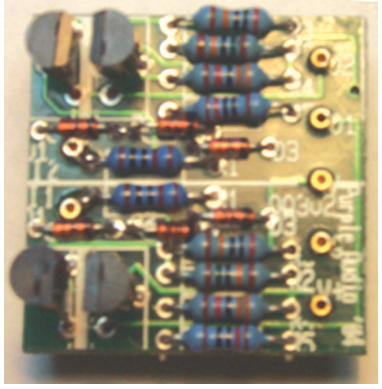


(*a*) (*b*)

Fig. 63.6

The output voltage e_0 is developed across emitter resistance R_E. When a signal e_s is applied at the input, the resulting a.c. emitter current i_e produces an output voltage $i_e R_E$ across R_E. This voltage opposes the signal voltage as it is in opposite direction to the signal. This voltage is fedback at the input of the amplifier. The feedback voltage is proportional to the emitter current *i.e.,* output current. So *the circuit provides negative current feedback*.

The output is taken from the emitter resistance R_E and the voltage variations across base-emitter junction (input) follow the emitter. So the circuit is named as an *emitter follower*. It is also called *common collector amplifier* because the transistor collector is at a.c. ground and common between input and output.



Emitter Follower.

The equivalent circuit, using *CE* hybrid model, of the emitter follower is given in Fig. 63.6 (*b*).

Voltage gain. The voltage gain *A* of the amplifier, without feedback, may be obtained by connecting the grounded end of e_s to *E*. Then,

$$e_s = (R_S + h_{ie})i_i \qquad \text{...(1)}$$

Output voltage $e_0 = h_{fe}i_iR_E$ ...(2)

Voltage gain without feedback

$$A = \frac{e_0}{e_S} = \frac{h_{fe}R_E}{R_S + h_{ie}} \qquad \text{...(3)}$$

Whole of the output voltage is fed to the input. So, feedback fraction $\beta = 1$.

i.e., $\beta e_0 = e_0$.

The voltage gain of the amplifier with negative feedback

$$A_f = \frac{A}{1+\beta A} = \frac{A}{1+A} \qquad (\because \beta = 1) \text{ ... (4)}$$

Thus the voltage gain of the emitter follower is always less than unity.

Substituting the value of A from Eq. (3) in Eq. (4), we get

$$A_f = \frac{\dfrac{h_{fe}R_E}{R_S + h_{ie}}}{1+\dfrac{h_{fe}R_E}{R_S + h_{ie}}} = \frac{h_{fe}R_E}{R_S + h_{ie} + h_{fe}R_E} \qquad \text{...(5)}$$

Input Impedance

Put $R_S = 0$ in Eq. (3). Then we get,

voltage gain without feedback

$$A = \frac{h_{fe}R_E}{h_{ie}} \qquad \text{...(6)}$$

From Fig. 63.6 (b), input impedance of the amplifier without feedback,

$$Z_i = h_{ie} \qquad \text{...(7)}$$

We know that input impedance of the amplifier with negative feedback,

$$Z_{if} = (1 + \beta A)Z_i$$

Substituting the value of A from Eq. (6) and of Z_i from Eq. (7) we get

$$Z_{if} = h_{ie}\left[1 + 1.\frac{h_{fe}R_E}{h_{ie}}\right] \qquad (\because \beta = 1)$$

or $Z_{if} = h_{ie} + h_{fe}R_E$ ...(8)

Thus the input impedance of emitter follower is considerably increased by negative feedback.

Output Impedance

The output impedance without feedback is the ratio of output voltage to output current.

Let e = open circuit output voltage

and i = short circuit output current *i.e.*, by connecting the grounded end of e_s to E in Fig. 63.6 (*b*).

$$e = \lim_{R_E \to \infty} e_0 = \lim_{R_E \to \infty} \frac{h_{fe}R_Ee_S}{R_S + h_{ie}} \qquad \text{[using Eq. (3)]}$$

$$i = \frac{h_{fe}e_S}{R_S + h_{ie}}$$

$$\therefore \quad Z_0 = \frac{e}{i} = \lim_{R_E \to \infty} R_E$$

Therefore, the output impedance without feedback is infinite.

Now, we know that output impedance of amplifier with negative feedback

$$Z_{of} = \frac{Z_0}{1+\beta A} = \frac{\lim\limits_{R_E \to \infty} R_E}{1+1\times\dfrac{h_{fe}R_E}{R_S+h_{ie}}} \qquad \text{[using Eq. (3)]}$$

$$= \lim_{R_E \to \infty} \frac{(R_S+h_{fe})\,R_E}{(R_S+h_{ie})+h_{fe}R_E}$$

$$= \lim_{R_E \to \infty} \frac{R_S+h_{fe}}{\dfrac{R_S+h_{ie}}{R_E}+h_{fe}}$$

$$= \frac{R_S+h_{fe}}{h_{fe}}\,.$$

Thus the output impedance of an emitter follower is considerably reduced by negative feedback.

63.4 Noise in Electronic Circuits

Electronic Noise Reduction Circuits.

The sh-h-h sound heard in a radio receiver tuned to a wavelength that has no station assigned to it or the snow-like appearance on a TV screen when the set is tuned to a weak station is due to the internal noise generated in the active and passive devices of the amplifiers. This noise places a lower limit on the amplifier's sensitivity, *i.e.*, on the amplitude of signal that can be detected and amplified without losing the signal into the noise.

The maximum amplification which can be used through cascading successive stages of amplification is limited by electrical *noise*, or erratic currents generated in the circuits or in the active devices. Gain sufficient to bring the internal noise up to the usual signal output level is thus the usable limit. Any input signal appreciably weaker than the noise will be covered up or masked and will be recoverable only through use of special techniques. Noise is introduced from a number of sources, some due to the inherent particle nature of electric current and some to passive circuit devices.

The possible effect of *noise* on a pulse signal is shown in Fig. 63.7.

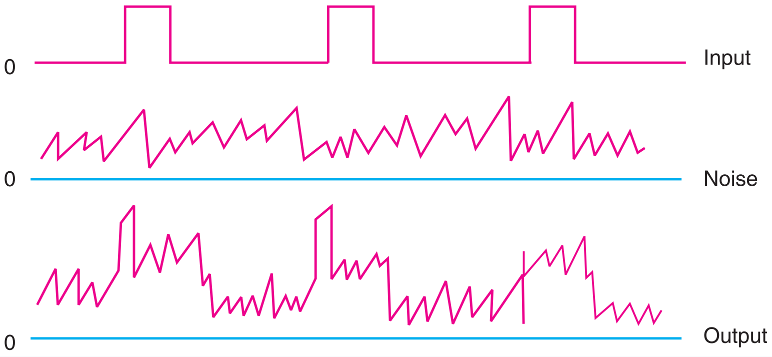


Fig. 63.7 Signals Plus Noise

It appears that an element of uncertainty is introduced into the time of arrival, the amplitude, and the width. It is usual to assume that the noise is “white” or that the noise power is uniformly distributed over the input-signal frequency spectrum. Reduction of the frequency band passed will then reduce the noise, but it may also reduce the information content of the signal.

Signal-to-Noise Ratio (S/N)

The S/N ratio is defined as the ratio of signal power to noise power at the same point in a circuit, *i.e.*,

$$\frac{S}{N} = \frac{P_s}{P_n} = \frac{V_s^2 / R}{V_n^2 / R} = \left(\frac{V_s}{V_n}\right)^2$$

Noise Figure

Noise is described by use of a *noise figure.*

The noise figure is defined as

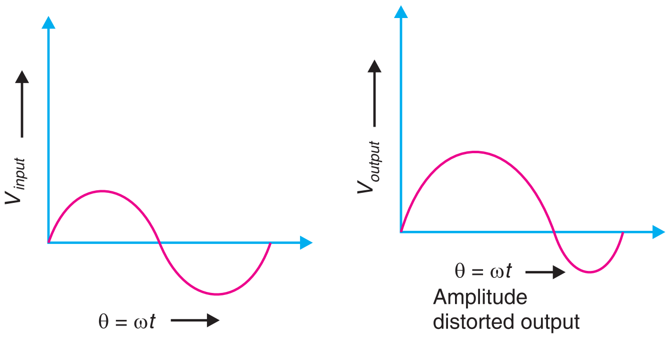
$$\text{N.F.} = \frac{P_{iS} / P_{iN}}{P_{oS} / P_{oN}} \text{ dB.}$$

Here, P_S/P_N is the signal-to-noise ratio at the indicated input and output of the circuit or device, indicating the degradation of the S/N ratio by the device.

63.5 Distortion in Amplifiers

An ideal amplifier must produce an amplified output voltage which must be an enlarged true copy of the input voltage, without any change in the wave-shape. Practical amplifiers, however, create some change in the wave-shape of the output. This is known as ‘distortion’.

There are three types of distortion.



(*a*) (*b*)

Fig. 63.8

***(i)* Amplitude (or Non-linear) Distortion.** The difference in the magnitudes of positive peak and negative peak of the amplified voltage is called amplitude distortion. Amplitude distorted waveform is shown in Fig. 63.8.

In this type of distortion, the output of the amplifier is not linearly related to the input. So the distortion is also called non-linear distortion. In large signal amplifiers, new frequencies are present in the output which are *not* present in the input signal. So the distortion is called harmonic distortion. Harmonic distortion can be reduced by using negative feedback circuits.

***(ii)* Frequency Distortion.** We know that the frequency response of an amplifier is not uniform. Whereas the voltage-gain is sufficiently uniform over the middle-frequency range, it sharply falls at low as well as at high frequencies. Therefore, if the input signal is a complex wave, certain frequency components are amplified more than others. The output wave is not a faithful reproduction of the signal. The distortion so produced is called ‘frequency distortion’.

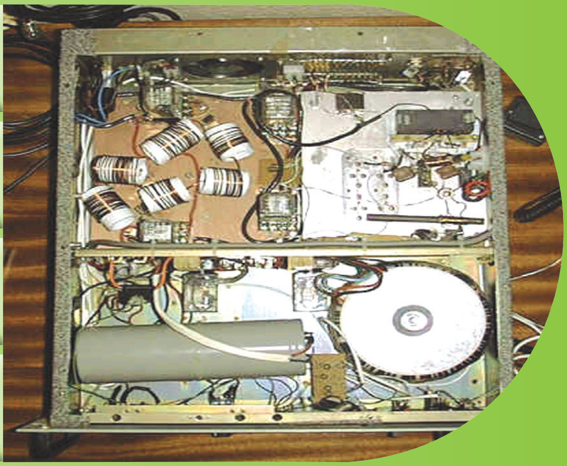
***(iii)* Phase Distortion.** In a practical amplifier, different frequencies may suffer different phase-changes. Therefore, relative phase of some components of a complex input wave may be shifted. This makes the shape of the output wave different from that of the signal. Special phase equalising circuits are employed to avoid phase distortion.

EXERCISE

1. Describe the working of an R – C coupled amplifier and explain its frequency response curve. (*Rohilkhand,* 88; *Meerut* 84, 86)
2. Give the circuit of R–C coupled amplifier. Obtain an expression for the voltage gain in mid-frequency range. (*Rohilkhand,* 90; *Nagpur,* 80, 81, 82, 88)
3. Discuss with a circuit diagram the working of a transistor R–C coupled amplifier. Explain why the voltage gain falls off at high and low frequencies of the signal. (*Mahatma Gandhi*, 1989)
4. Describe a transformer coupled amplifier. Discuss the frequency response curve.
5. Draw the diagram of emitter follower and its equivalent circuit. Derive expressions for its voltage gain, current gain, input impedance and output impedance. (*Nagarjuna,* 1993; *Andhra,* 95)

CHAPTER

64



FET AMPLIFIER

AT A GLANCE

64.1 Common Source FET Amplifier
64.2 Advantages of JFET over BJT
64.3 The Equivalent Circuit for the FET (Low frequency model)
64.4 JFET Source-Follower (Common-Drain) Amplifier
64.5 FET as Voltage-Variable Resistor (VVR)

64.1 Common Source FET Amplifier

Fig. 64.1 shows the common source amplifier circuit. V_{GG} is the gate bias battery. The ac signal to be amplified is applied between the gate and the source. The load resistance R_L is connected between the supply battery V_{DD} and the drain D. The output is taken across the load R_L or across the FET. Let the output ac voltage be v_{ds} for an input signal voltage v_{gs}.

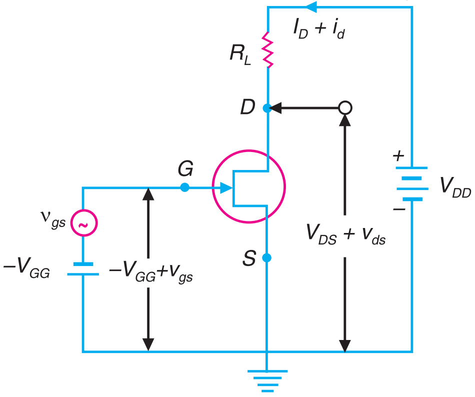


Fig. 64.1

Working. On applying an alternating signal, the fixed reverse bias of the gate changes. A small change in the reverse bias gate voltage, produces a large change in drain current. This fact makes FET capable of amplifying weak signal. The amplifying action of FET can be explained using its transfer characteristics (Fig. 64.2).

When the signal (v_{gs}) is not applied, the gate to source voltage is given by OA. The corresponding drain current is given by OF. When the signal is applied, during the positive half cycle of the input signal ($+ v_{gs}$), the reverse bias on the gate decreases ($- V_{GG} + v_{gs}$). The gate to source voltage decreases from OA to OC. Due to this, the channel width increases. Hence, drain current increases from OF to OG. During the negative half of the signal ($- v_{gs}$), the reverse voltage on the gate increases ($- V_{GG} - v_{gs}$) to OB. Due to this the channel width decreases. Hence the drain current decreases to OE. Thus, a small change in the gate to source voltage, produces a large change in drain current and hence large change in output voltage. Thus, FET works as an amplifier.

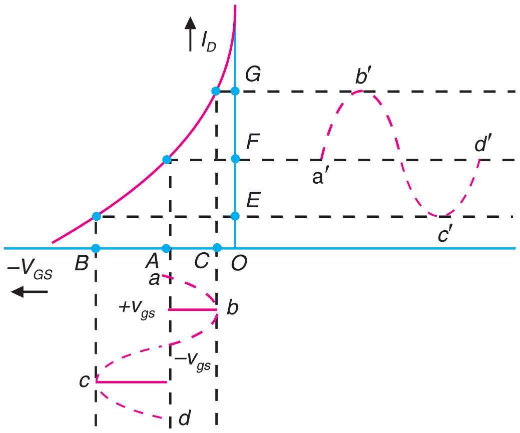


Fig. 64.2

When the signal voltage is positive, the gate becomes less negative with respect to the source. The drain current is enhanced, causing a large voltage drop across R_L which makes the drain terminal less positive with respect to the source. Since an increase of the gate voltage causes a decrease in the drain voltage, there is a phase shift of 180° between the input and output of the FET amplifier.

Expression for Voltage Gain

The a.c. voltage source equivalent circuit of FET amplifier is shown in Fig. 64.3. FET can be replaced by a voltage generator $-\mu v_{gs}$ in series with drain resistance r_{ds}. i_d is the ac current flowing through the load R_L. The output ac voltage is v_{ds} for an input signal voltage v_{gs}.

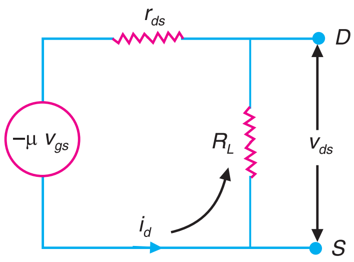


Fig. 64.3

Applying Kirchoff's voltage law to the circuit,

$$-\mu v_{gs} = i_d R_L + i_d r_{ds}$$

$$\therefore \quad i_d = \frac{-\mu v_{gs}}{r_{ds} + R_L}.$$

But

$$v_{ds} = i_d \,.\, R_L$$

$$i_d = \frac{v_{ds}}{R_L}$$

$$\therefore \quad \frac{v_{ds}}{R_L} = \frac{-\mu v_{gs}}{r_{ds} + R_L}$$

$$\frac{v_{ds}}{v_{gs}} = \frac{-\mu R_L}{r_{ds} + R_L}$$

Voltage gain

$$A_v = \frac{v_{ds}}{v_{gs}} = \frac{-\mu R_L}{r_{ds} + R_L}$$

This is the expression for voltage gain. The negative sign indicates that there is a phase difference of 180° between the input and the output voltages.

64.2 Advantages of JFET over BJT

	BJT	*FET*
1.	Bipolar	Unipolar
2.	Input impedance low	Input impedance high
3.	Current controlled	Voltage controlled
4.	Characterised by current gain	Characterised by transconductance
5.	Noise level not small	Noise level very small

Example 1. *A common-source FET amplifier has a load resistance* R_L = 500 kΩ. *If the a.c. drain resistance* (r_d) *and amplification factor* (μ) *of the FET are* 100 kΩ *and* 24, *respectively, calculate the voltage gain of the amplifier*.

Sol.

$$|A_v| = \frac{\mu R_L}{r_{ds} + R_L} = \frac{24 \times 500 \times 10^3}{100 \times 10^3 + 500 \times 10^3} = 20.$$

Example 2. *For a constant drain-to-source voltage if the gate-source voltage is changed from* 0 *to* –2 *V the corresponding change in drain current becomes* 2 *mA.Calculate the transconductance of the FET. If the a.c. drain resistance is* 100 kΩ, *calculate also the amplification factor of the FET.*

Sol. The transconductance g_m is given by

$$g_m = \left(\frac{\Delta I_D}{\Delta V_{GS}}\right)_{V_{DS}} = \frac{2\,mA}{2\,V} = 1\,\text{mA / V}.$$

Again,

$$\mu = g_m\, r_{ds} = 1 \times 10^{-3} \times 100 \times 10^3 = 100.$$

64.3 The Equivalent Circuit for the FET (Low frequency model)

Fig. 64.4 (*a*) shows the internal elements of a FET arranged as a two–port circuit. We can reduce it to a simpler equivalent form as in (*b*) by consideration of the relative magnitudes of the resistances and the reactances.

The input element r_{gs} is the reverse – biased junction resistance of a JFET or the silicon-dioxide insulation resistance of a MOSFET. These resistances exceed 10^8 ohms. We will consider the input of a FET as an open circuit.

The series circuit of C_c and r_c represents the capacitance between gate and channel and the series resistance of the channel. Being only a few picofarads and a few hundred ohms, the effect of this *RC* combination is negligible at low frequencies. So, they are eliminated from the equivalent circuit.

Capacitance C_{gd} is the gate-to-drain capacitance and includes the capacitance of the mounting. Being only 1 to 3pF, C_{gd} may be eliminated from the equivalent circuit.

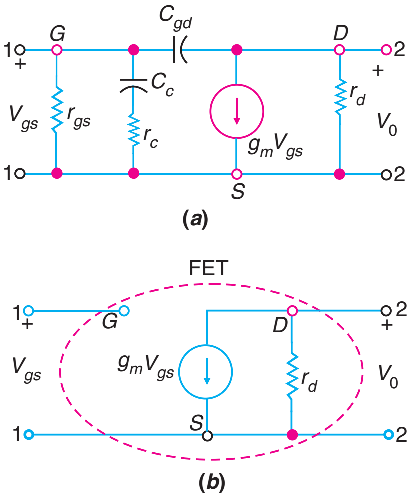


Fig. 64.4

(*a*) Internal FET circuit;
(*b*) Equivalent circuit for a FET.

FET ac equivalent circuit is shown in Fig. 64.4 (*b*). The control of I_d by V_{gs} is included as a current source $g_m V_{gs}$ connected from drain to source. The arrow points from drain to source to show a phase change of 180° between output and input voltages as will occur in actual operation.

Here gate to source voltage is now represented by V_{gs} (lower-case subscripts) to distinguish it from d.c. levels. In low frequency model as the gate is reversed biased, the gate current becomes zero. Therefore, the input impedance is represented by open circuit at input terminals. The output impedance is represented by the resistor r_d from drain to source. In situations, where r_d is sufficiently large in comparison to other circuit elements, this is ignored. Now, the equivalent circuit is simply a current source whose magnitude is controlled by V_{gs} and g_m-clearly a voltage controlled device.

64.4 JFET Source–Follower (Common–Drain) Amplifier

Fig. 64.5 shows the JFET source-follower configuration.

The input is applied between gate and source. The output is taken off the source terminal. When the d.c. supply is replaced by its short-circuit equivalent, the drain is grounded (hence, the terminology common-drain).

Fig. 64.6 shows the low frequency equivalent model for the common drain amplifier circuit shown in Fig. 64.5.

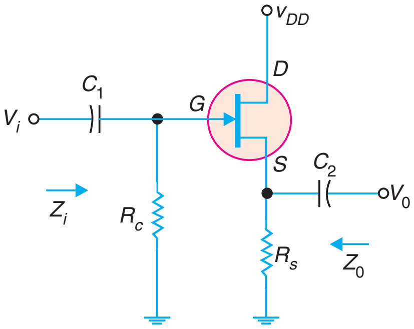


Fig. 64.5

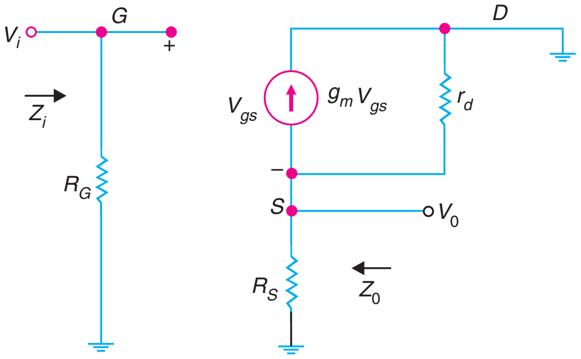


Fig. 64.6

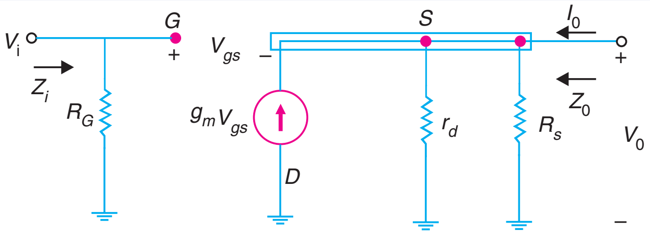


Fig. 64.7

Fig. 64.6 can be simplifed as shown in Fig. 64.7.

Input Impedance Z_i : Fig. 64.7 reveals that Z_i is defined by

$$Z_i = R_G \qquad \text{...(1)}$$

Output Impedance Z_0 : Fig. 64.8 shows the circuit for Z_0 determination. The output inpedance of the circuit is found by making the independent signal source, V_i, equal zero. Setting $V_i = 0$ V will result in the gate terminal being connected directly to ground.

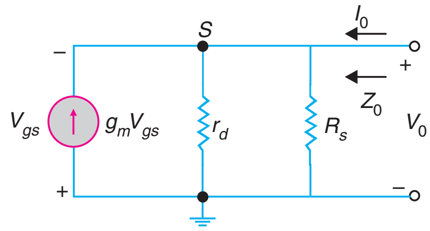


Fig. 64.8

V_{gs} and V_0 are across the same parallel network.

$\therefore \qquad V_0 = -V_{gs} \qquad$ ...(2)

Applying Kirchhoff's current law at node S,

$$I_0 + g_m V_{gs} = I_{rd} + I_{Rs} = \frac{V_0}{r_d} + \frac{V_0}{R_S}$$

$$I_0 = V_0\left[\frac{1}{r_d} + \frac{1}{R_S}\right] - g_m V_{gs}$$

$$= V_0\left[\frac{1}{r_d} + \frac{1}{R_S}\right] - g_m[-V_0] = V_0\left[\frac{1}{r_d} + \frac{1}{R_S} + g_m\right]$$

$$Z_0 = \frac{V_0}{I_0} = \frac{V_0}{V_0\left[\frac{1}{r_d} + \frac{1}{R_S} + g_m\right]} = \frac{1}{\frac{1}{r_d} + \frac{1}{R_S} + g_m}$$

$$= \frac{1}{\frac{1}{r_d} + \frac{1}{R_S} + \frac{1}{1/g_m}}$$

$\therefore \qquad Z_0 = r_d \| R_S \| 1/g_m \qquad$ ...(3)

For $r_d \geq 10\, R_S$, $\qquad Z_0 \cong R_S \| 1/g_m \qquad$ ...(4)

Expression for Voltage Gain

The output voltage V_0 is determined by

$$V_0 = g_m V_{gs} (r_d \| R_S) \qquad \text{...(5)}$$

Applying KVL around the perimeter of the network of Fig. 64.7,

$$V_i = V_{gs} + V_0$$

or $\qquad V_{gs} = V_i - V_0$

Eq. (5) becomes $\qquad V_0 = g_m(V_i - V_0)(r_d \| R_S)$

or $\qquad V_0 = g_m V_i (r_d \| R_S) - g_m V_0 (r_d \| R_S)$

or $\qquad V_0[1 + g_m(r_d \| R_S)] = g_m V_i (r_d \| R_S)$

$\therefore \qquad A_V = \dfrac{V_0}{V_i} = \dfrac{g_m (r_d \| R_S)}{1 + g_m (r_d \| R_S)} \qquad$ ...(6)

If $r_d \geq 10\, R_S$, $\qquad A_V = \dfrac{V_0}{V_i} \cong \dfrac{g_m R_S}{1 + g_m R_S} \qquad$ ...(7)

If $g_m\, R_S >> 1 \qquad A_V = 1$, but it is always less than one.

We observe that common drain circuit does not provide voltage gain.

Phase Relationship: Since A_v of Eq. (6) is a positive quantity, V_0 and V_i, are in phase for the JFET source-follower configuration.

64.5 FET as Voltage-Variable Resistor (VVR)

Principle. JFET is operated in the constant-current portion of its output characteristics *i.e.*, in saturation region for the linear applications. FET can also be used in the region before pinch-off where V_{DS} is small. FET when used in the region before pinch-off, it works as variable resistance device *i.e.*, the channel resistance r_d is controlled by the gate bias voltage (V_{GS}).

In JFET, the drain-to-source conductance $g_d = \dfrac{I_D}{V_{DS}}$ for small values of V_{DS}. It may also be expressed as,

$$g_d = g_{d0}\left[1 - \left(\frac{V_{GS}}{V_P}\right)^{\frac{1}{2}}\right]$$

where g_{d0} is the value of drain conductance when the bias voltage V_{GS} is zero.

The variation of the r_d with V_{GS} can be closely approximated by the empirical expression,

$$r_d = \frac{r_0}{1 - KV_{GS}}$$

where r_0 = drain resistance at zero gate bias, and K = a constant, dependent upon FET type.

Thus, small signal FET drain resistance r_d varies with applied gate voltage V_{GS} and FET acts like a variable passive resistor.

FET finds wide applications where VVR property is useful. For example, the VVR can be used in *Automatic Gain Control* (AGC) circuit of a multistage amplifier.

The VVR is used to vary the voltage gain of a multistage amplifier. If the signal is low then voltage gain of the stages can be increased and when the signal becomes high, the gain can be reduced automatically. In this way, the general level of amplification is maintained fairly constant.

Circuit. Fig. 64.9 shows the circuit arrangement of AGC amplifier using the FET as *Voltage Variable Resistor* (*VVR*) or *Voltage Dependent Resistor* (*VDR*).

Working. The input signal v_i is amplified by amplifier. It is then rectified and filtered to produce a d.c. voltage proportional to output signal level. This voltage is applied to the gate of the FET so that the a.c. resistance between drain and source changes.

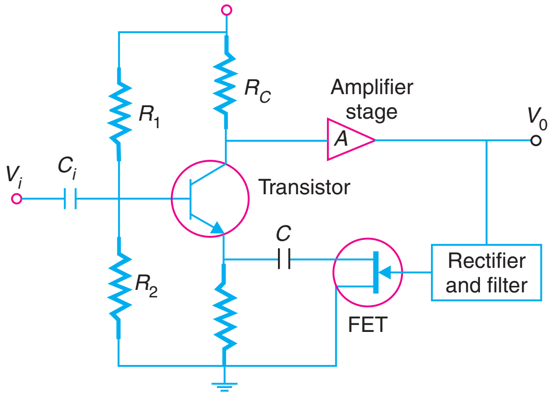


Fig. 64.9

The capacitor C isolates the transistor from FET so that the bias conditions of transistor are not affected. So, when output increases, V_{GS} also increases and R_{DS} changes so that the gain of the transistor decreases. Thus automatically the gain is controlled.

EXERCISE

1. Draw the circuit diagram of a basic FET amplifier and explain its operation. Draw the ac voltage source equivalent circuit of the FET amplifier.
2. A common-source FET amplifier uses a load resistance $R_L = 200\ k\Omega$. If the a.c. drain resistance and the transconductance of the FET are 100 $k\Omega$ and 0.15 mA/V, respectively, calculate the magnitude of the voltage gain of the amplifier. (**Ans.** 10)

CHAPTER

65

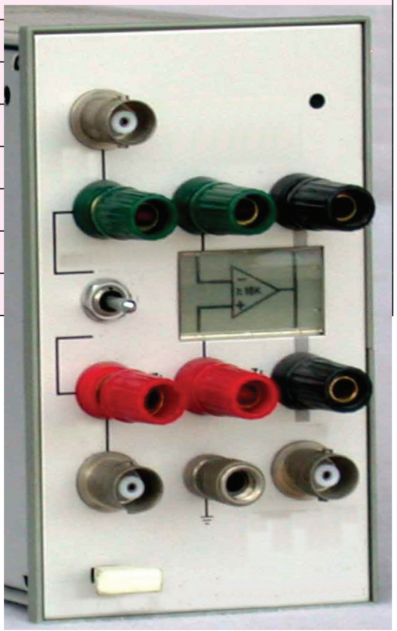
OPERATIONAL AMPLIFIERS

AT A GLANCE

65.1 Introduction
65.2 Characteristics of an OP AMP
65.3 The Common Mode Rejection Ratio (CMRR)
65.4 The Slew Rate
65.5 Inverting Amplifier
65.6 Adder or Summing Amplifier
65.7 Op-Amp as Integrator
65.8 Op-Amp as Differentiator
65.9 Difference Amplifier or Subtractor
65.10 Logarithmic Amplifier
65.11 Op-Amp as Comparator
65.12 Non-inverting Operational Amplifier
65.13 Voltage Follower

65.1 Introduction

An operational amplifier is a high gain direct coupled amplifier with high input impedance and low output impedance to which feedback is added to regulate overall response.



Operational Amplifier.

Circuit Symbol of an *OP AMP*

Fig. 65.1 shows the circuit symbol of an *OP* AMP.

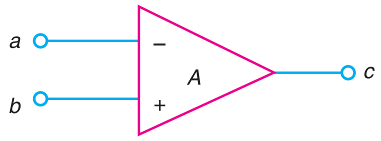


Fig. 65.1

Terminals *a* and *b* are the input terminals. The terminal *c* is the output terminal. Terminal *a* (marked '–') is called the *inverting input terminal*. The negative sign indicates that a signal applied at the terminal *a* will appear at the terminal *c* with a polarity opposite to that at the terminal *a*. Terminal *b* (marked '+') is the *noninverting input terminal*. This means that the output signal at *c* is always of the same polarity as that of the signal applied at the terminal *b*.

The output signal voltage is proportional to the *difference* between the two signal voltages applied at the two input terminals.

$$V_0 = AV_1 - AV_2 = A(V_1 - V_2).$$

Here, voltages V_1 and V_2 are measured with respect to ground (Fig. 65.2).

A is the differential voltage gain. The quantity *A* is a real constant approaching infinity in the ideal case for all frequencies.

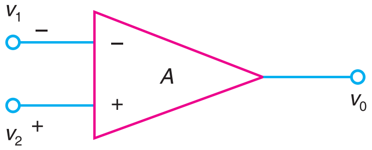


Fig. 65.2

65.2 Characteristics of an OP AMP

An ideal operational amplifier has the following characteristics.

1. Infinite input impedance *i.e.*, $Z_i = \infty$
2. Zero output impedance $Z_0 = 0$
3. Infinite voltage gain $A = -\infty$

   Even a feeble differential input signal gets amplified to a large extent.
4. Infinite bandwidth $B_w = \infty$

   The voltage gain of the amplifier is constant at all frequencies of the input signals.
5. Perfect balance $V_0 = 0$ when $V_1 = V_2$.

   Any signal, common to both the inputs, is rejected at the output.
6. Zero drift, *i.e.,* characteristics do not change with temperature.
7. Common mode rejection ratio should tend to ∞.
8. Slew rate should tend to ∞.

65.3 The Common Mode Rejection Ratio (CMRR)

It is defined as the ratio between the differential gain, A_d to the common mode gain, A_c, i.*e.,*

$$CMRR = \rho = \frac{A_d}{A_c} = \frac{\text{Differential–mode gain}}{\text{Common–mode gain}}.$$

Alternately, *CMR* may be expressed in decibels as

$$CMR = 20 \log CMRR = 20 \log A_d - 20 \log A_c.$$

CMRR is infinity for a differential amplifier.

65.4 The Slew Rate

It is defined as the maximum rate of change in output voltage per unit of time

$$SR = \left.\frac{dV_0}{dt}\right|_{\max}.$$

It is expressed in volts per microsecond.

Slew rate indicates how fast the output of an op-amp can change in response to changes in the input frequency.

EXAMPLE. *The op-amp 741 has a slew rate of 0.5 V/μs. Its implication is explained by considering Fig. 65.3.*

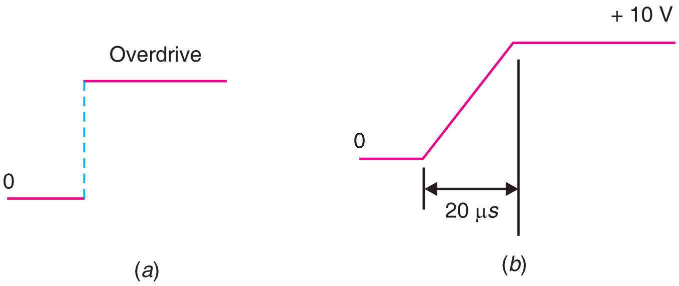


Fig. 65.3

If we overdrive a 741 with a large step input (Fig. 65.3 *a*), the output slews as shown in Fig. 65.3 *b*. It takes 20 microseconds for the output voltage to change from 0 to 10 V. It is impossible for the 741 to change faster than this.

65.5 Inverting Amplifier

If only one input is applied to the inverting input terminal, then it is called inverting amplifier. Fig 65.4 shows the circuit of an inverting amplifier.

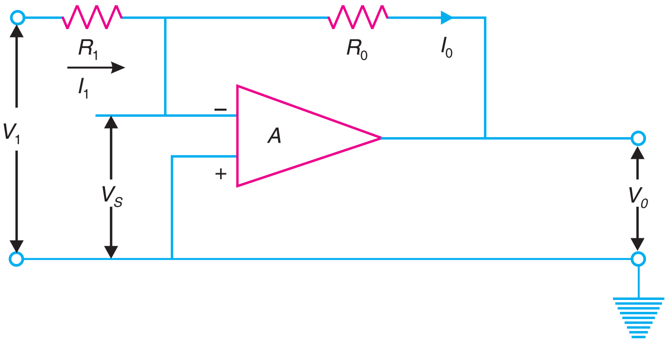
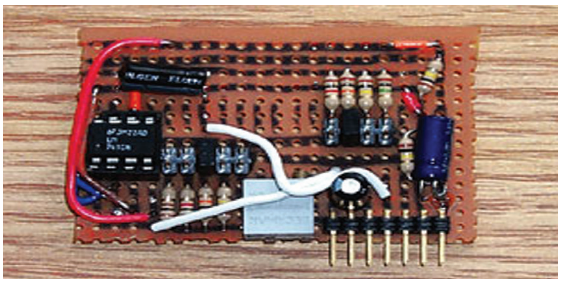


Fig. 65.4

The differential input resistance R_1 is infinite for an ideal *OP-AMP*. But for a practical amplifier it must be much larger than output resistance R_0. So we can assume that $I_0 = I_1$. It means that an extremely negligible amount of current flows in the *OP-AMP*. We may write

$$I_1 = \frac{V_1 - V_s}{R_1} \qquad \text{...(1)}$$

and $$I_0 = \frac{V_s - V_0}{R_0} \qquad \text{...(2)}$$



Inverting Amplifier.

Neglecting the current entering into the *OP–AMP*, we may write

$$I_1 = I_0$$

$$\frac{V_1 - V_s}{R_1} = \frac{V_s - V_0}{R_0}$$

On rearranging $$\frac{V_0}{R_0} = -\frac{V_1}{R_1} + \frac{V_s}{R_1} + \frac{V_s}{R_0} \quad ...(3)$$

Again $$V_0 = -AV_s \text{ or } V_s = -\frac{V_0}{A}$$

$$\frac{V_0}{R_0} = -\frac{V_1}{R_1} - \frac{V_0}{A}\left(\frac{1}{R_0} + \frac{1}{R_1}\right) \quad ...(4)$$

On the right hand side of Eq. (4), the second term may be neglected compared with the first term. Therefore, Eq. (4) can be written as

$$\frac{V_0}{R_0} = -\frac{V_1}{R_1}$$

or voltage gain, $$\frac{V_0}{V_1} = -\frac{R_0}{R_1} \quad ...(5)$$

Thus the resulting amplification is entirely determined by feedback return *i.e.,* by resistors R_0 and R_1.

OP- AMP APPLICATIONS

65.6 Adder or Summing Amplifier

An adder is a circuit whose output is proportional to the algebraic sum of the input voltages.

Consider an inverting amplifier with 3 inputs at the inverting terminal. (Fig. 65.5)

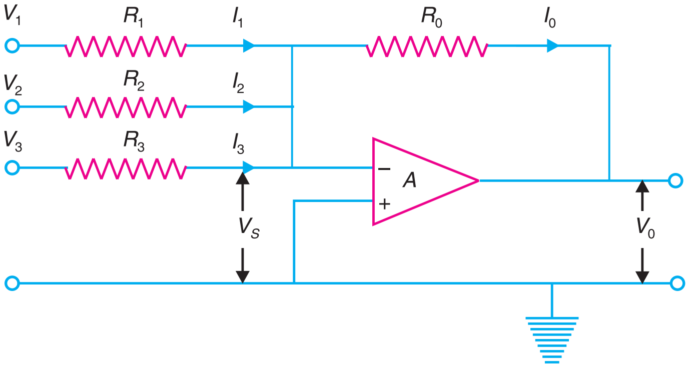


Fig. 65.5

The inverting terminal is virtually grounded by the feedback resistor R_0.

So the sum of the currents through R_1, R_2, R_3 is equal to the current through R_0.

Summing Amplifier.

$$I_0 = I_1 + I_2 + I_3 \qquad \text{...(1)}$$

$$\frac{V_s - V_0}{R_0} = \frac{V_1 - V_s}{R_1} + \frac{V_2 - V_s}{R_2} + \frac{V_3 - V_s}{R_3}$$

or
$$-\frac{V_0}{R_0} = \frac{V_1}{R_1} + \frac{V_2}{R_2} + \frac{V_3}{R_3} - V_s\left(\frac{1}{R_0} + \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}\right) \qquad \text{... (2)}$$

We have
$$V_0 = -AV_s.$$

$$-\frac{V_0}{R_0} = \left(\frac{V_1}{R_1} + \frac{V_2}{R_2} + \frac{V_3}{R_3}\right) + \frac{V_0}{A}\left(\frac{1}{R_0} + \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}\right) \qquad \text{... (3)}$$

Since the value of A is very large, the second term on the R.H.S. is neglected as compared with the first term. Eq. (3) becomes

$$-\frac{V_0}{R_0} = \frac{V_1}{R_1} + \frac{V_2}{R_2} + \frac{V_3}{R_3} \qquad \text{...(4)}$$

If
$$R_0 = R_1 = R_2 = R_3, \text{ we get}$$

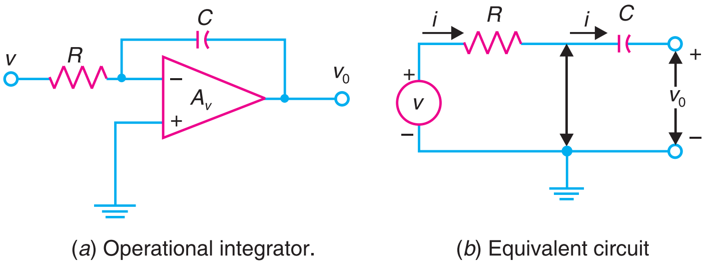
$$V_0 = -(V_1 + V_2 + V_3) \qquad \text{...(5)}$$

output voltage = – (sum of the input voltages).

65.7 Op-Amp as Integrator

The circuit shown in Fig. 65.6 produces an output voltage that is proportional to the time integral of the input voltage. Hence, this circuit is called an integrator.

The input signal source of voltage $v(t)$ is connected to the inverting input terminal through resistance R. The negative feedback is given using a capacitor C. The non-inverting terminal is earth-connected [Fig. 65.6(a)]



(*a*) Operational integrator. (*b*) Equivalent circuit

Fig. 65.6

In Fig. 65.6 *b*, the double-headed arrow represents a virtual ground. Hence $i = v/R$.

$$v_0 = -\frac{1}{C}\int i\,dt = -\frac{1}{RC}\int v\,dt \qquad \text{...(1)}$$

The amplifier therefore provides an output voltage proportional to the integral of the input voltage.

Case I. If the input voltage is a constant, $v = V$, then the output will be a ramp, $V_0 = -Vt/RC$. Such an integrator makes an excellent sweep circuit for a cathode-ray-tube oscilloscope, and is called a *Miller integrator*, or *Miller sweep*.

Case II. If the input voltage is a square wave, then the output will be a triangle wave (Fig. 65.7).

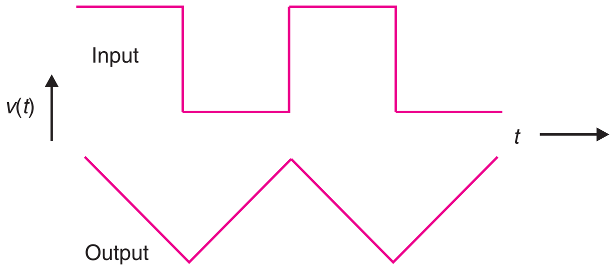
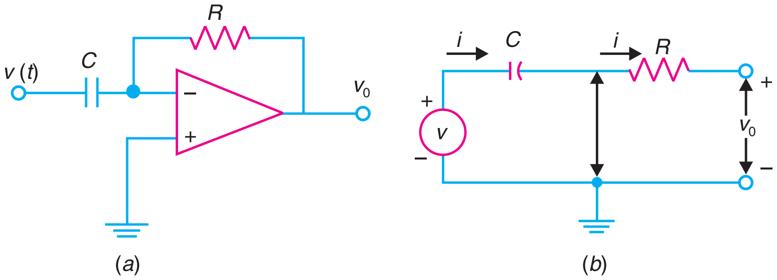


Fig. 65.7

The integrating circuits are used to solve differential equations in analog computers.

65.8 Op-Amp as Differentiator

Fig. 65.8 (*a*) shows a circuit that performs the operation of differentiation.



(*a*) (*b*)

Fig. 65.8

The input signal source of voltage *v* (*t*) is connected to the inverting input terminal through a capacitor *C*. The non-inverting input terminal is earth-connected. Negative feedback is given through a resistance *R*.

Let *v* (*t*) be the signal voltage given as the input, which drives varying current through the capacitance *C*.

We see from the equivalent circuit of Fig. 65.8 (*b*) that

$$i = C\frac{dv(t)}{dt} \qquad ...(1)$$

$$v_0 = -Ri = -RC\frac{d}{dt}v(t) \qquad ...(2)$$

Hence, the output voltage is proportional to the differential of the input voltage.

65.9 Difference Amplifier or Subtractor

The function of a subtracter is to provide an output proportional to or equal to the difference of two input signals.

We have to apply the inputs at the inverting and non-inverting terminals (Fig. 65.9).

Both inverting and non-inverting operations take place simultaneously.

The output can be derived using the superposition theorem.

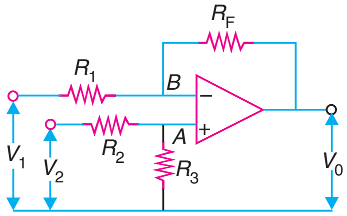


Fig. 65.9

Thus the resultant output,

$$v_0 = v_{01} + v_{02} \qquad \text{...(1)}$$

Here, v_{01} and v_{02} are the outputs due to v_1 and v_2 inputs respectively.

Output v_{01} : The output v_{01} due to v_1 alone with the other input earthed is the same as that of inverting circuit.

$$v_{01} = -(R_F/R_1)v_1 \qquad \text{...(2)}$$

Output v_{02} : The output v_{02} due to v_2 alone with the other input earthed is the same as that of non-inverting circuit.

$$v_{02} = (1 + R_F/R_1)v_A$$

Here, $v_A = R_3\, v_2 /(R_2 + R_3)$

$$v_{02} = \left(1+\frac{R_F}{R_1}\right)\frac{R_3}{R_2 + R_3}v_2 \qquad \text{...(3)}$$

The resultant $v_0 = -\dfrac{R_F}{R_1}v_1 + \left(1+\dfrac{R_F}{R_1}\right)\left(\dfrac{R_3}{R_2 + R_3}\right)v_2$ ...(4)

For $R_2 = R_1$ and $R_3 = R_F$, we get

$$v_0 = -\frac{R_F}{R_1}v_1 + \left(1+\frac{R_F}{R_1}\right)\left(\frac{R_F}{R_1 + R_F}\right)v_2$$

$$= (R_F/R_1)(v_2 - v_1) \qquad \text{...(5)}$$

Further, if $R_F = R_1$, then,

$$v_0 = (v_2 - v_1) \qquad \text{...(6)}$$

Output voltage = difference of the two input voltages

65.10 Logarithmic Amplifier

Principle. A logarithmic amplifier has an output voltage which is proportional to the logarithm of the input voltage, *i.e.*,

$$v_0 \propto \log_e v_i.$$

The linear OPAMP can be combined with a nonlinear element such as a diode or transistor to achieve this. The output gets greatly compressed. Therefore the response of a meter across the output will be like that of a decibel meter.

(*i*) Fig. 65.10 shows the circuit of a logarithmic amplifier using a diode.

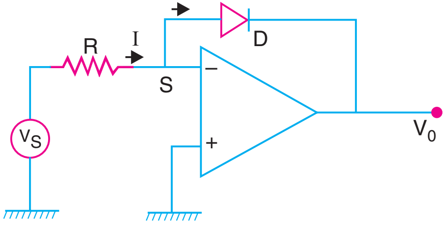


Fig. 65.10

The branch point *S* acts as the virtual ground. When the input voltage source is switched on, current flowing through the resistance *R* flows through the diode *D*.

Let I = forward current through the diode and

V = forward voltage across the diode.

The relation between the current and voltage is

$$I = I_0(e^{qV/\eta kT} - 1) \qquad \text{...(1)}$$

Here, q is the electronic charge, I_0 is the reverse saturation current of the diode, k is the Boltzman constant, T is the kelvin temperature of the diode. $\eta = 1$ for germanium and 2 for silicon diode.

Let $\dfrac{kT}{q} = x$. Then, $I = I_0\,(e^{V/\eta x} - 1) = I_0 e^{V/\eta x}$ $\left(\because \dfrac{V}{\eta x} >> 1\right)$

By virtual ground property of the input terminal ,
current through R = current through the diode,

$$\frac{V_S}{R} = \text{I}$$

$$\therefore \qquad (V_S/R) = I_0 e^{V/\eta x}$$

$$\frac{V_S}{RI_0} = e^{V/\eta x}$$

Taking logarithm, $\log_e\left(\frac{V_S}{RI_0}\right) = \frac{V}{\eta x}$

$$\therefore \qquad V = \eta x \log_e(V_S/RI_0) \qquad ...(2)$$

The forward voltage across the diode $V = 0 - V_0$

or $\qquad V_0 = -V$

The output voltage $V_0 = -\eta x \log_e\left(\frac{V_S}{RI_0}\right)$ ...(3)

All the quantities on the right hand side of the above equation (except V_S) are constants at constant temperature. So, the output voltage V_0 is proportional to the logarithm of the input voltage V_S, and the circuit works as logarithmic amplifier.

(*ii*) Though a diode can be used for realising a log-amplifier, the grounded base transistor is preferred because the exponential relation between current and voltage extends over a much wider voltage range. A typical circuit is shown in Fig. 65.11.

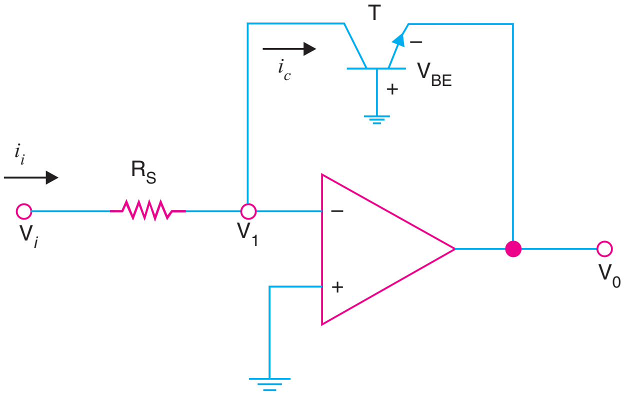


Fig. 65.11

Since $v_1 = 0$ and the base is grounded, $v_{CB} = 0$.

The collector current $i_C = \beta i_B$...(1)

Here, the base current is given by

$$i_B = I_S \exp(qv_{BE}/kT)$$

which is the same as the diode current of the forward -biased emitter-base *p-n* junction. Hence

$$i_C = \beta I_S \exp(qv_{BE}/kT) \qquad ...(2)$$

$$= K \exp(qv_{BE}/kT) \qquad ...(3)$$

$$\therefore \qquad v_{BE} = \frac{kT}{q} \ln\left(\frac{i_c}{K}\right) \qquad ...(4)$$

Since the emitter of the transistor is connected to the output of the OPAMP,

$$v_0 = v_{EB} = -v_{BE}$$

Also, $\qquad i_C = i_i = \frac{v_i}{R_S}$ ...(5)

Hence $\qquad v_0 = -\frac{kT}{q} \ln\left(\frac{v_i}{KR_S}\right)$ ...(6)

Thus v_0 is proportional to log v_i.

65.11 Op-Amp as Comparator

A comparator is a circuit with two inputs. It is used to compare a voltage in one input with a fixed reference voltage at the other input.

We can examine the operation of a comparator using a 741 op-amp, as shown in Fig. 65.12.

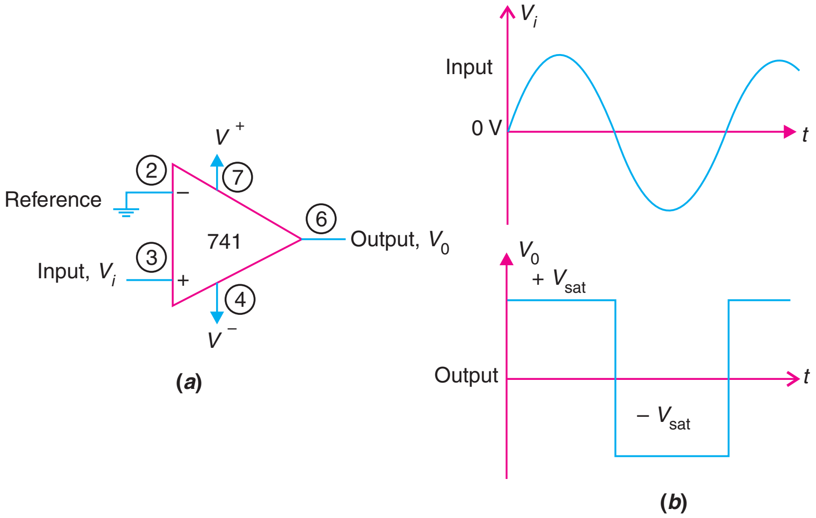


Fig. 65.12

With reference input (at pin 2) set to 0V, a sinusoidal signal applied to the noninverting input (pin 3) will cause the output to switch between its two output states, as shown in Fig. 65.12 (*b*). The input V_i going even a fraction of a millivolt above the 0-V reference level will be amplified by the very high voltage gain (typically over 100,000). So the output rises to its positive output saturation level and remains there while the input stays above $V_{ref} = 0\ V$.

When the input drops just below the 0-V reference level, the output is driven to its lower saturation level and stays there while the input remains below $V_{ref} = 0\ V$.

The input signal is linear. The output signal is digital.

In general use, the reference level need not be 0 V but can be any desired positive or negative voltage. Also, the reference voltage may be connected to either plus or minus input and the input signal then applied to the other input.

65.12 Non-inverting Operational Amplifier

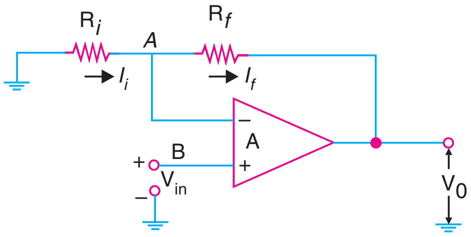


Fig. 65.13

Fig. 65.13 shows the circuit diagram of an ideal op- amp in the non-inverting mode. The input signal is applied to the non- inverting input terminal (+). The feed back is applied to the inverting input terminal (–) through R_f. The resistors R_f and R_i form the feed back voltage divider circuit.

To find the voltage gain

The node B is at potential V_{in}. Hence the potential of point A is the same as that of B which is V_{in}.

Applying KCL,

$$\frac{0 - V_{in}}{R_i} = \frac{V_{in} - V_0}{R_f}$$

$$\therefore \qquad -\frac{V_{in}}{R_f} - \frac{V_{in}}{R_i} = -\frac{V_0}{R_f}$$

$$\therefore \qquad V_{in}\left[\frac{1}{R_i} + \frac{1}{R_f}\right] = \frac{V_0}{R_f}$$

$$V_0 = V_{in}\left[\frac{R_f}{R_i} + 1\right]$$

$$\therefore \qquad \frac{V_0}{V_{in}} = 1 + \frac{R_f}{R_i}$$

The output voltage is in phase with the input voltage (Fig. 65.14).

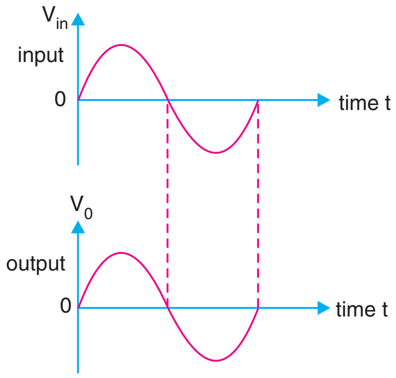


Fig. 65.14

65.13 Voltage Follower

A circuit in which the output voltage follows the input voltage is called voltage follower circuit.

Fig. 65.15 shows the voltage follower circuit using op-amp.

$$V_A = V_B = V_{in} \qquad \text{...(1)}$$

The node B is at potential V_{in}. Now node A is also at the same potential as B, *i.e.*, V_{in}.

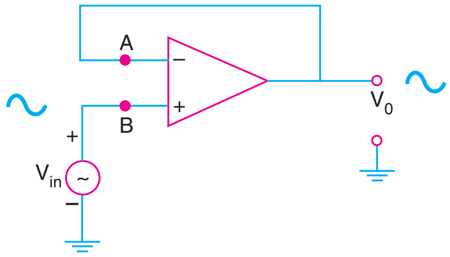


Fig. 65.15

Node A is connected directly to the output.

$$\therefore \qquad V_0 = V_A \qquad \text{...(2)}$$

From Eqs. (1) and (2),

$$V_0 = V_{in} \qquad \text{...(3)}$$

i.e., output voltage = input voltage both in magnitude and phase.

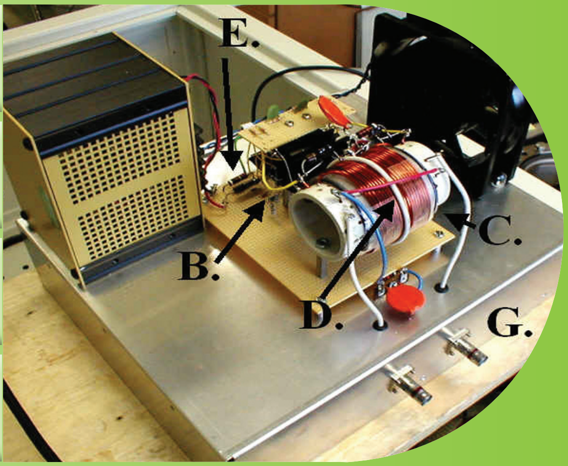
The voltage gain of the circuit becomes unity. The circuit is then referred to as a *voltage follower* or a *unity-gain buffer*. As the circuit has a high input impedance and a low output impedance, it can be used as an impedance matching device between a high-impedance source and a low-impedance load.

A voltage buffer circuit provides a means of isolating an input signal from a load by using a stage having unity voltage gain, with no phase or polarity inversion, and acting as an ideal circuit with very high input impedance and low output impedance.

EXERCISE

1. What are the different characteristics of an ideal operational amplifier ? (*Osmania,* 1993; *Andhra,* 1993; *S.K.U.*, 1994)
2. Define CMRR. (*Andhra,* 1993, 95)
3. Define the term slew rate. (*Andhra,* 1993, 95)
4. Draw the circuit diagram of an inverting amplifier and explain its operation. Derive an expression for its voltage gain. (*Andhra,* 1995)
5. Explain how an Op-Amp can be used as (*i*) an adder (*ii*) an integrator (*iii*) a differentiator. Obtain expression for output in each case.

CHAPTER 66



OSCILLATORS

AT A GLANCE			
66.1	Hartley Oscillator	**66.2**	Colpitt's Oscillator
66.3	Wien-Bridge Oscillator	**66.4**	Phase Shift Oscillator
66.5	Crystal Oscillator		
66.6	Wien Bridge Oscillator		

66.1 Hartley Oscillator

Figure 66.1 shows the circuit of a shunt-fed Hartley oscillator. The resistors R_1, R_2 and R_E and the supply voltage V_{cc} establish the d.c. operating point of the transistor. The frequency determining

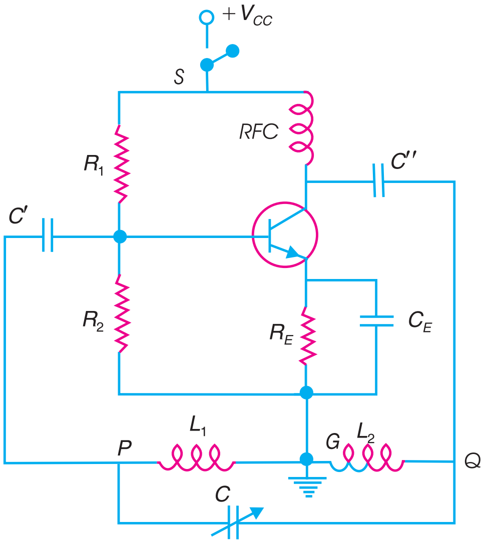


Fig. 66.1

Hartley Oscillator.

network is made up of the variable capacitor C and the inductors L_1 and L_2. The coil L_1 is inductively coupled to coil L_2 and the combination forms an autotransformer. The capacitor C'' blocks dc and provides an ac path from the collector to the tank circuit. Radio frequency choke RFC provides d.c. load for the collector and also prevents a.c. from reaching the d.c. supply V_{cc}. Capacitor C' and resistor R_2 produce self bias voltage. Capacitor C_E provides a.c. ground and prevents any signal degeneration. C_E also provides temperature stabilization. The input to the amplifier appears to the base through the capacitor C'. The output is available at the collector through the capacitor C'' with respect to the ground.

Working. When the switch S is closed, the collector current starts increasing and charges the capacitor C. When this capacitor is fully charged, it discharges through coils L_1 and L_2, setting up damped harmonic oscillations in the tank circuit. The oscillatory current in the tank circuit produces an a.c. voltage across L_1 which is applied to the base emitter junction of the transistor and appears in amplified form in the collector circuit. Feedback of energy from output (collector emitter circuit) to input (base emitter circuit) is accomplished through autotransformer action. Thus the energy is fedback into the tank circuit by means of mutual inductance between L_1 and L_2. This energy supplied to the tank circuit overcomes the losses occurring in it. Consequently the oscillations are sustained in the circuit.

Because of earth connection at G, the potentials at Q and P are 180° out of phase. Thus the phase difference between the voltage across L_1 and that across L_2 is always 180°. A further phase shift of 180° is introduced between the input and output voltage by the transistor itself. Thus the total phase shift becomes 360°, thereby making the feedback positive which is essential for oscillations. Consequently, continuous undamped oscillations are obtained.

Expression for frequency

Fig. 66.2 shows the equivalent circuit of the oscillator. Z_1, Z_2 and Z_3 represent the inductive or capacitive reactances of oscillator. Here the feedback source of e.m.f. $h_{re}\ V_0$ has been omitted because reverse voltage h_{re} of a transistor is negligible. The output resistance $1/h_{0e}$ in parallel with Z_2 is very large. Therefore, it has also been omitted.

The input resistance h_{ie} of the transistor and impedance Z_1 are in parallel. Their equivalent impedance Z′ is given by

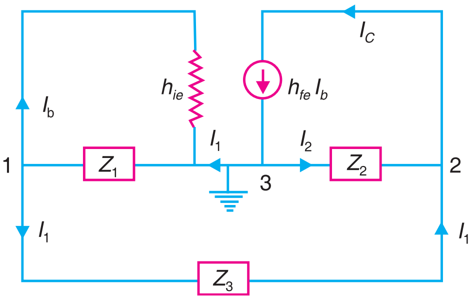


Fig. 66.2

$$\frac{1}{Z'} = \frac{1}{Z_1} + \frac{1}{h_{ie}} = \frac{Z_1 + h_{ie}}{Z_1\,h_{ie}}$$

or
$$Z' = \frac{Z_1\ h_{ie}}{Z_1 + h_{ie}} \qquad ...(1)$$

This equivalent impedance Z' is in series with Z_3. This combination acts in parallel with Z_2. Thus the *load impedance* (Z_1) between the output terminals 2 and 3 is given by

$$\frac{1}{Z_L} = \frac{1}{Z_2} + \frac{1}{Z' + Z_3}$$

$$\frac{1}{Z_L} = \frac{1}{Z_2} + \frac{1}{\dfrac{Z_1 h_{ie}}{Z_1 + h_{ie}} + Z_3}$$

$$= \frac{1}{Z_2} + \frac{Z_1 + h_{ie}}{Z_1 h_{ie} + Z_1 Z_3 + Z_3 h_{ie}}$$

$$= \frac{1}{Z_2} + \frac{Z_1 + h_{ie}}{Z_1 Z_3 + h_{ie} (Z_1 + Z_3)}$$

$$= \frac{Z_1 Z_3 + h_{ie} (Z_1 + Z_3) + Z_2 (Z_1 + h_{ie})}{Z_2 [Z_1 Z_3 + h_{ie} (Z_1 + Z_3)]}$$

or
$$Z_L = \frac{Z_2 [Z_1 Z_3 + h_{ie} (Z_1 + Z_3)]}{h_{ie} (Z_1 + Z_2 + Z_3) + Z_1 Z_2 + Z_1 Z_3} \qquad ...(2)$$

The *output voltage* between terminals 3 and 2 in terms of current I_1 is

$$V_0 = -I_1(Z' + Z_3) = -I_1\left(\frac{Z_1 h_{ie}}{Z_1 + h_{ie}} + Z_3\right) = -I_1\left(\frac{Z_1 h_{ie} + Z_1 Z_3 + Z_3 h_{ie}}{Z_1 + h_{ie}}\right)$$

$$V_0 = -I_1\left[\frac{Z_1 Z_3 + h_{ie} (Z_1 + Z_3)}{Z_1 + h_{ie}}\right] \qquad ...(3)$$

Voltage feedback to the input terminals 3 and 1 is given by

$$V_{fb} = -I_1 Z' = -I_1\left(\frac{Z_1 h_{ie}}{Z_1 + h_{ie}}\right) \qquad ...(4)$$

Therefore, the feedback fraction β is given by

$$\beta = \frac{V_{fb}}{V_0} = I_1\left(\frac{Z_1 h_{ie}}{Z_1 + h_{ie}}\right)\left[\frac{Z_1 + h_{ie}}{Z_1 Z_3 + h_{ie} (Z_1 + Z_3)}\right]\frac{1}{I_1}$$

$$\beta = \frac{Z_1 h_{ie}}{Z_1 Z_3 + h_{ie} (Z_1 + Z_3)} \qquad ...(5)$$

Now from Barkhausen criterion, the condition for the maintenance of oscillation is

$$\beta A_v = 1 \qquad ...(6)$$

Here, A_v is the voltage gain without feedback. For a *CE* amplifier,

$$A_v = -\frac{h_{fe}}{h_{ie}} Z_L \qquad ...(7)$$

Substituting the value of β from Eq. (5) and A_v from Eq. (7) in Eq. (6), we get

$$\left(-\frac{h_{fe} Z_L}{h_{ie}}\right)\left[\frac{Z_1 h_{ie}}{Z_1 Z_3 + h_{ie} (Z_1 + Z_3)}\right] = 1$$

or
$$\left[\frac{h_{fe} Z_2 \{Z_1 Z_3 + h_{ie} (Z_1 + Z_3)\}}{h_{ie} (Z_1 + Z_2 + Z_3) + Z_1 Z_2 + Z_1 Z_3}\right]\left[\frac{Z_1}{Z_1 Z_3 + h_{ie} (Z_1 + Z_3)}\right] = -1$$

or $$\frac{h_{fe} Z_1 Z_2}{h_{ie}(Z_1 + Z_2 + Z_3) + Z_1 Z_2 + Z_1 Z_3} = -1$$

or $$h_{ie}(Z_1 + Z_2 + Z_3) + Z_1 Z_2 + Z_1 Z_3 = -h_{fe} Z_1 Z_2$$

or $$h_{ie}(Z_1 + Z_2 + Z_3) + Z_1 Z_2 (1 + h_{fe}) + Z_1 Z_3 = 0 \qquad ...(8)$$

Now, we apply this general equation to the Hartley oscillator.

Let M be the mutual inductance between the two coils of self inductance L_1 and L_2. C is the capacitance in the tank circuit. Then

$$Z_1 = j\omega L_1 + j\omega M$$

$$Z_2 = j\omega L_2 + j\omega M$$

and $$Z_3 = \frac{1}{j\omega C} = -\frac{j}{\omega C}$$

Therefore, Eq. (8) gives

$$h_{ie}\left[(j\omega L_1 + j\omega M) + (j\omega L_2 + j\omega M) - \frac{j}{\omega C}\right]$$

$$+ (j\omega L_1 + j\omega M)(j\omega L_2 + j\omega M)(1 + h_{fe})$$

$$+ (j\omega L_1 + j\omega M)\left(-\frac{j}{\omega C}\right) = 0$$

or $$j\omega h_{ie}\left[L_1 + L_2 + 2M - \frac{1}{\omega^2 C}\right]$$

$$- \omega^2 (L_1 + M)(L_2 + M)(1 + h_{fe}) + \frac{L_1 + M}{C} = 0$$

or $$j\omega h_{ie}\left[L_1 + L_2 + 2M - \frac{1}{\omega^2 C}\right]$$

$$- \omega^2 (L_1 + M)\left[(L_2 + M)(1 + h_{fe}) - \frac{1}{\omega^2 C}\right] = 0 \qquad ...(9)$$

Equating the imaginary part on both sides of Eq. (9), we get

$$L_1 + L_2 + 2M - \frac{1}{\omega^2 C} = 0$$

or $$\omega^2 C = \frac{1}{L_1 + L_2 + 2M} \qquad ...(10)$$

or $$\omega = \frac{1}{\sqrt{\{(L_1 + L_2 + 2M)C\}}}$$

Frequency of oscillation,

$$f = \frac{\omega}{2\pi} = \frac{1}{2\pi\sqrt{\{(L_1 + L_2 + 2M)C\}}} \qquad ...(11)$$

The *condition for maintenance of oscillations* is found by equating the real part of Eq. (9) to zero. Thus

$$(L_2 + M)(1 + h_{fe}) - \frac{1}{\omega^2 C} = 0$$

$$\text{or} \qquad 1 + h_{fe} = \frac{1}{\omega^2 C (L_2 + M)}$$

Substituting the value of $\omega^2 C$ from Eq. (10), we get

$$1 + h_{fe} = \frac{L_1 + L_2 + 2M}{L_2 + M}$$

$$= \frac{(L_1 + M) + (L_2 + M)}{L_2 + M} = 1 + \frac{L_1 + M}{L_2 + M}$$

$$\text{or} \qquad h_{fe} = \frac{L_1 + M}{L_2 + M} \qquad \text{...(12)}$$

Eq. (12) gives the condition for sustained oscillation.

66.2 Colpitt's Oscillator

Fig. 66.3 shows the circuit of a Colpitt's Oscillator. Two capacitors C_1 and C_2 are placed across a common inductance L and the centre of the two capacitors is tapped. The resistors R_1 and R_2 form a voltage divider across the supply voltage V_{CC} and provide necessary bias condition for the circuit along with R_E. The function of capacitor C'' is to block d.c. and provide an a.c. path from collector to the tank circuit. C_E is a bypass capacitor. Radio frequency choke RFC provides the necessary d.c. load resistance for collector and also prevents a.c. signal from entering the d.c. supply V_{CC}. C'' conveys feedback from the collector to base circuit.

Working. When the switch S is closed, the collector current starts increasing. Capacitors C_1 and C_2 are charged. These capacitors discharge through coil L setting up damped harmonic oscillations in the tank circuit. The oscillatory current in the tank circuit produces an a.c. voltage across C_1. The voltage across C_1 is applied to base-emitter junction of the transistor and appears in amplified form in the collector circuit. It overcomes the losses occurring in the tank circuit.

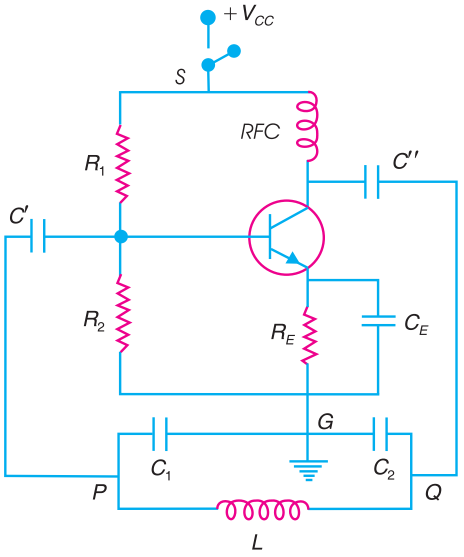


Fig. 66.3

The energy supplied to the tank circuit is in phase with the generated oscillations. The feedback voltage (voltage across the capacitor C_1) is 180° out of phase with the output voltage (voltage across the capacitor C_2). It is because, the terminal G is earthed which will be at a zero potential. When P is at a positive potential with respect to G, Q will be at a negative potential with respect to G at that instant. Therefore, points P and Q are 180° out of phase. A further phase shift of 180° between the output and input voltage is produced by common emitter transistor. Thus the total phase shift is 360°. Thus the phase shift condition for sustained oscillation is satisfied and we get continuous undamped oscillations.

Expression for Frequency

The frequency and condition of oscillations in Colpitt's oscillator can be obtained by applying the *general equation of oscillator* [Eq. (8) in Hartley Oscillator]. That is,

$$h_{ie}(Z_1 + Z_2 + Z_3) + Z_1 Z_2 (1 + h_{fe}) + Z_1 Z_3 = 0.$$

Here, $$Z_1 = \frac{1}{j\omega C_1} = -\frac{j}{\omega C_1}$$

$$Z_2 = \frac{1}{j\omega C_2} = -\frac{j}{\omega C_2}$$

and $$Z_3 = j\omega L.$$

Substituting these values in the general equation, we get

$$h_{ie}\left(-\frac{j}{\omega C_1} - \frac{j}{\omega C_2} + j\omega L\right) + \left(-\frac{j}{\omega C_1}\right)\left(-\frac{j}{\omega C_2}\right)(1 + h_{fe}) + \left(-\frac{j}{\omega C_1}\right)j\omega L = 0$$

or $$-jh_{ie}\left(\frac{1}{\omega C_1} + \frac{1}{\omega C_2} - \omega L\right) - \frac{1 + h_{fe}}{\omega^2 C_1 C_2} + \frac{L}{C_1} = 0$$

or $$\left(\frac{1 + h_{fe}}{\omega^2 C_1 C_2} - \frac{L}{C_1}\right) + jh_{ie}\left(\frac{1}{\omega C_1} + \frac{1}{\omega C_2} - \omega L\right) = 0 \qquad \text{...(1)}$$

Equating the imaginary parts on both sides of Eq. (1), we get

$$\frac{1}{\omega C_1} + \frac{1}{\omega C_2} - \omega L = 0$$

or $$\frac{C_1 + C_2}{\omega C_1 C_2} - \omega L = 0$$

or $$\omega^2 L C_1 C_2 = C_1 + C_2 \qquad \text{...(2)}$$

or $$\omega = \sqrt{\left(\frac{C_1 + C_2}{LC_1 C_2}\right)}$$

Frequency of oscillation

$$f = \frac{\omega}{2\pi} = \frac{1}{2\pi}\sqrt{\left(\frac{C_1 + C_2}{LC_1 C_2}\right)} \qquad \text{...(3)}$$

Equating the real parts on both sides of Eq. (1), we get

$$\frac{1 + h_{fe}}{\omega^2 C_1 C_2} - \frac{L}{C_1} = 0 \quad \text{or } 1 + {}_{fe} = \frac{LC_1 C_2}{\ }$$

Substituting the value of $\omega^2 LC_1 C_2$ from Eq. (2), we get

$$1 + h_{fe} = \frac{C_1 + C_2}{C_1} = 1 + \frac{C_2}{C_1}$$

or $$h_{fe} = \frac{C_2}{C_1} \qquad \text{...(4)}$$

This is the condition for sustained oscillations.

66.3 Wien-Bridge Oscillator

Figure 66.4 shows the Wien bridge network. The frequency of oscillations is determined by the series element R_1 C_1 and parallel element R_2 C_2. The upper arm consists of resistance R and capacitance C in series. The lower arm contains a parallel combination of equal resistance and capacitance.

The ratio of output voltage (V_0) of the network to the input voltage (V_i) is

$$\frac{V_0}{V_i} = \frac{\text{Impedance of parallel combination}}{\text{Total impedance}}$$

$$\frac{V_0}{V_i} = \frac{-\dfrac{(R \cdot j/\omega C)}{R - j/\omega C}}{R - \dfrac{j}{\omega C} - \dfrac{(R \cdot j/\omega C)}{R - j/\omega C}}$$

$$= \frac{-jR/\omega C}{\left(R - \dfrac{j}{\omega C}\right)^2 - \dfrac{jR}{\omega C}} = \frac{-jR/\omega C}{\left(R^2 - \dfrac{1}{\omega^2 C^2} - \dfrac{3jR}{\omega C}\right)} \qquad \text{...(1)}$$

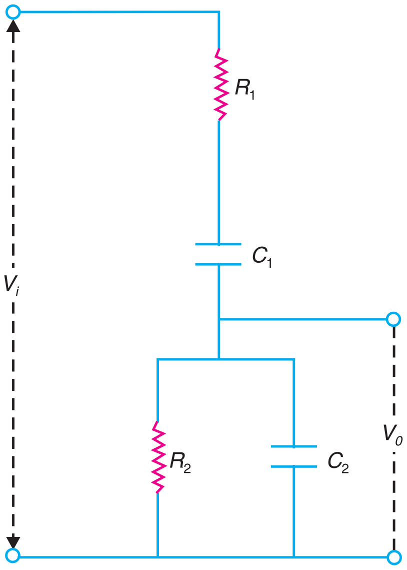


Fig. 66.4

The phase shift will be zero if imaginary term vanishes, *i.e.*,

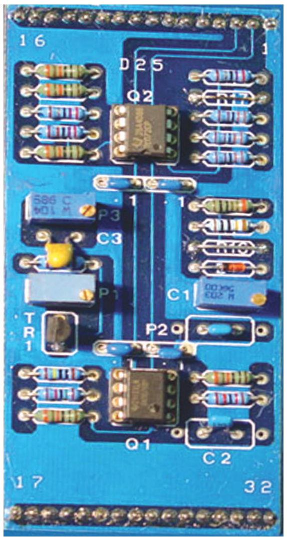
$$R^2 - \frac{1}{\omega^2 C^2} = 0 \qquad \text{or} \qquad \omega = \frac{1}{RC}.$$

Frequency of oscillation

$$f_0 = \frac{\omega}{2\pi} = \frac{1}{2\pi RC} \qquad \text{...(2)}$$

From Eq. (1),
$$\frac{}{V} = \frac{-jR/\omega C}{3jR/\ \ C} = \frac{}{3} \qquad \text{...(3)}$$

Thus the oscillations will be sustained if the amplifier has a gain just exceeding 3.



Wien-Bridge Oscillator.

Circuit Diagram. Figure 66.5 shows the circuit diagram of Wien bridge oscillator. It is basically a two stage transistor amplifier. The first stage is a common-emitter amplifier. The second one is a common collector amplifier. The output of the second stage is fedback to the input of the first stage through feedback network consisting of R_1 C_1 in series and R_2 C_2 in parallel. If

$R_1 = R_2 = R$ and $C_1 = C_2 = C$, then

$$f = \frac{1}{2\pi RC}$$

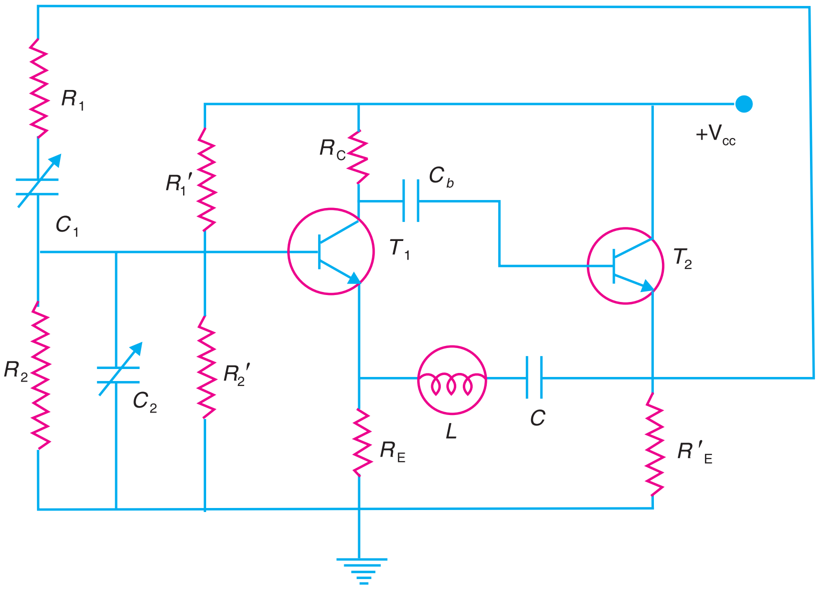


Fig. 66.5

Continuous variation of frequency is accomplished by varying simultaneously the two capacitors (ganged variable air capacitors). Changes in frequency range are accomplished by switching in different values for the two identical resistors *R*.

Working. An input signal applied at the base of transistor T_1 appears in the amplified form across collector resistor R_C but it simultaneously suffers a phase change of 180°. A further change of 180° is introduced by transistor T_2. Thus the total phase shift is either 360° or 0°. Hence the feedback will be positive and circuit self-oscillatory.

The *R-C* element of Wien-bridge provides a negative feedback which prevents oscillations at all frequencies. Thus the oscillator becomes sensitive to a signal of only one particular frequency. Thus Wien-bridge oscillator has good frequency stability.

L is a temperature sensitive tungsten lamp. It is included in the circuit to limit the amplitude of oscillations. If there is an increase in amplitude of oscillations, the lamp resistance increases reducing the positive feedback and hence the amplitude of oscillations.

When bridge oscillator finds wide use in variable audio frequency generators. It is a low frequency, low distortion sine wave generator . It gives constant output.

66.4 Phase Shift Oscillator

Principle. Consider an *RC* network, connected in series with a signal generator (Fig. 66.6*a*). Let a sinusoidal voltage $V_i = E_0 \sin \omega t$ of frequency $f = \dfrac{\omega}{2\pi}$ be applied to the network. The alternating voltage V_0 across *R* leads the applied voltage V_i by an angle ϕ, given by

$$\tan \phi = \frac{1}{\omega CR} = \frac{1}{2\pi fCR}.$$

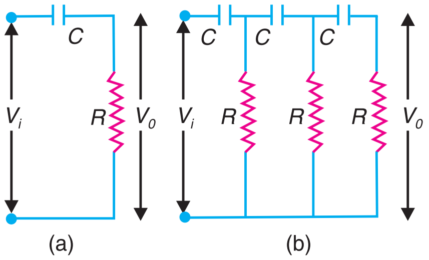


Fig. 66.6

For a given frequency f, by adjusting the values of C and R, ϕ can be made equal to 60°.

In a phase shift oscillator, a ladder network consisting of three RC sections is used (Fig. 66.6 b). Each section will produce a phase shift of 60°.

Consequently a total phase shift of 180° is produced between the input and output voltages *i.e.*, V_0 leads V_i by 180°.

Phase Shift Oscillator.

Circuit diagram. Figure 66.7 shows a transistor phase shift oscillator in CE mode. R_L is the collector load resistor. C_E is the emitter bypass capacitor. The resistors R_1, R_2 and R_E provide the necessary bias conditions for the circuit.

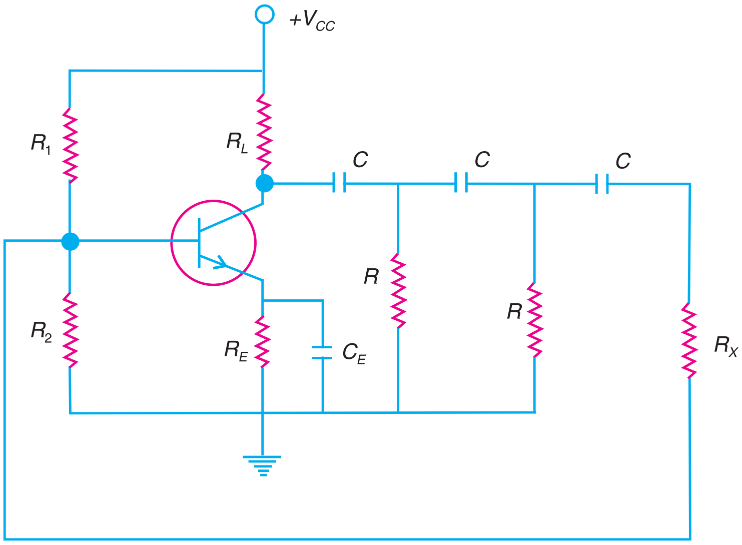


Fig. 66.7

For feedback of energy from collector to base, a ladder network consisting of three identical *RC* sections is connected between collector and base. The last section contains a resistance R_x such that

$$R_x + h_{ie} = R$$

where h_{ie} is the input resistance of the transistor. Since this resistance is connected with the base of the transistor, the input resistance h_{ie} of the transistor is added to it to give a total resistance of *R*.

Working. The circuit is set into oscillations by any random variation caused in the base current by the minor variation in the oscillator power supply V_{CC} or by the noise inherent in the transistor. This variation in the base current is amplified in the collector circuit and then fedback to the *RC* network. The network produces a phase shift of 180° between its input and output voltages. An additional phase shift is produced by the common emitter transistor. Consequently the total phase shift around the entire loop is 360° or 0° and the circuit acts as an oscillator. The oscillator produces sine wave output voltage continuously.

The frequency of oscillation is given by

$$f = \frac{1}{2\pi C\sqrt{(4RR_L + 6R^2)}}\,.$$

Expression for frequency of oscillation

The equivalent circuit of phase shift oscillator is shown in Fig. 66.8.

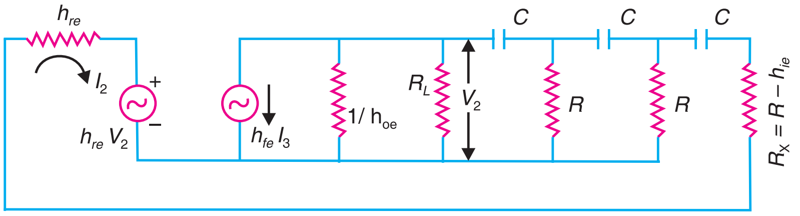


Fig. 66.8

In this circuit $1/h_{0e}$ can be neglected because it is much larger than R_L. Also, $h_{re}\ V_2$ can be neglected because h_{re} of a transistor is very small. We can replace the current source $h_{fe}\ I_3$ by an equivalent Thevenin voltage source of generated voltage $h_{fe}\ I_3\ R_L$. So the reduced equivalent circuit is shown in Fig. 66.9.

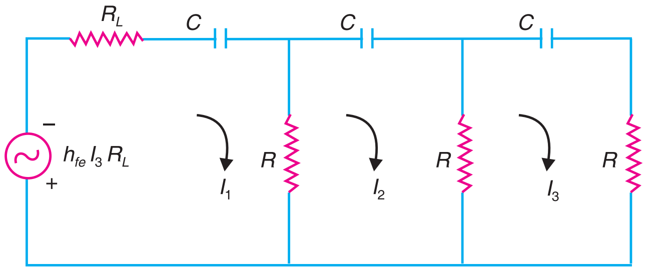


Fig. 66.9

Applying Kirchhoff's voltage laws to the three meshes, we get

$$(R_L + R - jX_c)\, I_1 - RI_2 + h_{fe}\, R_L\, I_3 = 0 \qquad ...(1)$$

$$-RI_1 + (2R - jX_c)\, I_2 - RI_3 = 0 \qquad ...(2)$$

and

$$0 \times I_1 - RI_2 + (2R - jX_c)\, I_3 = 0 \qquad ...(3)$$

Here, $X_c = \dfrac{1}{\omega C}$ is the capacitive reactance.

The currents I_1, I_2 and I_3 are non-zero. So the determinant of the coefficients of I_1, I_2 and I_3 must be zero.

$$\begin{vmatrix} R_L + R - jX_c & (-R) & h_{fe} R_L \\ -R & (2R - jX_c) & (-R) \\ 0 & (-R) & (2R - jX_c) \end{vmatrix} = 0 \qquad \text{...(4)}$$

$$(R_L + R - jX_c)\,(3R^2 - 4RjX_c - X_c^2) - R^2\,(2R - jX_c) + h_{fe}\,R_L\,R^2 = 0 \qquad \text{...(5)}$$

Equating the imaginary parts of Eq. (5), we get

$$-\,4RX_c\,(R_L + R) - X_c\,(3R^2 - X_c^2) + R^2X_c = 0$$

or $$-\,4RR_L - 4R^2 - 3R^2 + X_c^2 + R^2 = 0$$

or $$X_c^2 = 4RR_L + 6R^2 \qquad \text{...(6)}$$

or $$\frac{1}{\omega^2 C^2} = 4RR_L + 6R^2 \quad \text{or} \quad \omega^2 = \frac{1}{C^2\,(4RR_L + 6R^2)}$$

$\therefore$ $$f = \frac{\omega}{2\pi} = \frac{1}{2\pi C\sqrt{(4RR_L + 6R^2)}} \qquad \text{...(7)}$$

Condition for Maintenance of Oscillations

Equating the real parts of Eq. (5), we get

$$(R_L + R)\,(3R^2 - X_c^2) - 4RX_c^2 - 2R^3 + h_{fe}\,R_L\,R^2 = 0$$

or $$3R^2R_L + 3R^3 - R_LX_c^2 - RX_c^2 - 4RX_c^2 - 2R^3 + h_{fe}\,R_L\,R^2 = 0$$

or $$3R^2R_L + R^3 - X_c^2\,(5R + R_L) + h_{fe}\,R_L\,R^2 = 0$$

From Eq. (6), $X_c^2 = 4RR_L + 6R^2$

$\therefore$ $$3R^2R_L + R^3 - (4RR_L + 6R^2)\,(5R + R_L) + h_{fe}\,R_L\,R^2 = 0$$

or $$h_{fe}\,R_L\,R^2 = 23R^2\,R_L + 29R^3 + 4RR_L^2$$

$\therefore$ $$h_{fe} = 23 + 29\frac{R}{R_L} + 4\frac{R_L}{R} \qquad \text{...(9)}$$

Usually, $R = R_L$. Hence *the condition of sustained oscillations* is

$$h_{fe} = 56.$$

Thus for sustained oscillations in phase shift oscillator, the minimum value of the forward current gain ratio of the transistor should be 56.

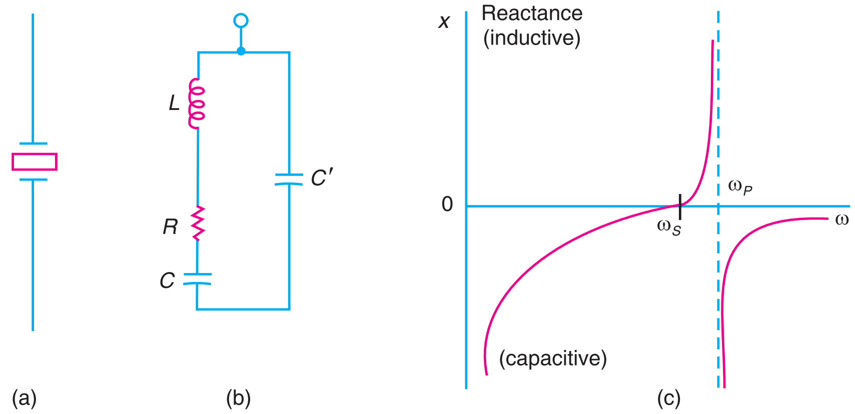
Variable-frequency Operation. The phase-shift oscillator is particularly suited to the range of frequencies from several hertz to several hundred kilohertz, and so includes the range of audio frequencies. The frequency of oscillation may be varied by changing any of the impedance elements in the phase-shifting network. For variations of frequency over a large range, the three capacitors are usually varied simultaneously. The phase-shift oscillator is operated in class *A* in order to keep distortion to a minimum.

66.5 Crystal Oscillator

(c) 1998 JAMES P. HAWKINS - WA2WHV

Crystal Oscillator.

Principle. When a suitably cut piezo-electric crystal is subjected to an appropriate alternating voltage, it vibrates mechanically. The amplitude of the mechanical oscillations becomes maximum when the frequency of the applied a.c. voltage is equal to the natural frequency of the crystal. Fig. 66.10 (*a*), shows the symbol of a piezoelectric crystal. The electrical equivalent circuit of a crystal is indicated in Fig. 66.10 (*b*). The inductor *L*, capacitor *C*, and resistor *R* are the analogs of the mass, the compliance (the reciprocal of the spring constant), and the viscous-damping factor of the mechanical system. *C′* represents the electrostatic capacitance between electrodes with the crystal as a dielectric.



(a) (b) (c)

Fig. 66.10

We neglect the resistance *R*. The impedance of the crystal is a reactance *jX* whose dependence upon frequency is given by

$$jX = -\frac{j}{\omega C'}\,\frac{\omega^2 - \omega_s^2}{\omega^2 - \omega_p^2}\,.$$

Here, $\omega_s^2 = \dfrac{1}{LC}$ or $f_s = \dfrac{1}{2\pi\sqrt{LC}}$.

It is called the *series resonant frequency* (the zero impedance frequency). At f_s, the crystal acts as a series resonant circuit.

$$\omega_p^2 = \frac{1}{L}\left(\frac{1}{C} + \frac{1}{C'}\right) \qquad \text{or} \qquad f_p = \frac{1}{2\pi\sqrt{L\left(\dfrac{CC'}{C + C'}\right)}}\,.$$

It is called the *parallel resonant frequency* (the infinite impedance frequency). At f_p, the crystal acts as a parallel resonant circuit.

Since $C' >> C$ then $f_p \approx f_s$. For $\omega_s < \omega < \omega_p$, the reactance is inductive, and outside this range it is capacitive, as indicated in Fig. 66.10 (*c*).

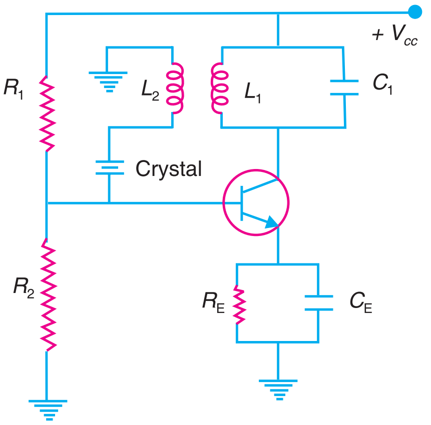


Fig. 66.11

Circuit Diagram. Figure 66.11 shows the transistor crystal oscillator. The battery V_{CC}, resistors R_1, R_2 and R_E and capacitor C_E provide the required dc bias for the transistor. A tank circuit L_1 C_1 is connected in the collector circuit. The coil L_2 is inductively coupled to coil L_1. One end of L_2 is connected to the base through the crystal and the other end is earthed. The resonance frequency of L_1 C_1 circuit is made nearly equal to the natural frequency of the crystal.

Working. When the dc source V_{CC} is switched on, a transient current flows through the tank circuit and the capacitor C_1 is charged. When it discharges, oscillations are set up in the tank circuit. The voltage across L_1 is fed to coil L_2 by mutual inductance. The voltages across L_1 and L_2 produce a phase difference of 180° between the output (collector) and input (base) circuits. A phase shift of 180° is produced by the transistor. The total phase shift between the output and input voltages is 360°. So the feedback is positive. This is the condition for oscillations. Since the crystal is connected in the base circuit, the frequency of oscillations is controlled by the crystal. Consequently, the L_1 C_1 circuit vibrates at the natural frequency of the crystal.

Advantages. 1. Crystal oscillator has a high order of frequency stability. The circuit provides a constant frequency because the vibration of crystal is independent of temperature, transistor parameters etc. It mainly depends on the nature in which the crystal is cut and on its thickness.

2. Q factor of the crystal is given by $Q = \frac{L\omega}{R}$ and it is of the order of 10^6. For greater frequency stability, Q factor should be high.

66.6 Wien Bridge Oscillator

Circuit Diagram A practical oscillator circuit uses an op-amp and *RC* bridge circuit, with the oscillator frequency set by the *R* and *C* components. Fig. 66.12 shows the Wien bridge oscillator circuit using op-amp amplifier.

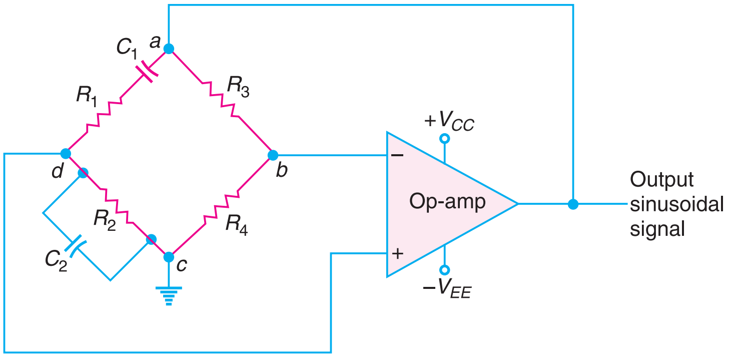


Fig. 66.12

(*i*) Resistors R_1, R_2 and capacitors C_1, C_2 form the frequency-adjustment elements.

(*ii*) Resistors R_3 and R_4 form part of the feedback path.

(*iii*) The op-amp output is connected as the bridge input at points *a* and *c*.

(*iv*) The bridge circuit output at points *b* and *d* is the input to the op-amp.

Frequency of Oscillation

Neglecting loading effects of the op-amp input and output impedances, the analysis of the bridge circuit results in

$$\frac{R_3}{R_4} = \frac{R_1}{R_2} + \frac{C_2}{C_1} \qquad \text{...(1)}$$

and $$f_0 = \frac{1}{2\pi\sqrt{R_1 C_1 R_2 C_2}} \qquad \text{...(2)}$$

If $R_1 = R_2 = R$ and $C_1 = C_2 = C$, the oscillator frequency is

$$f_0 = \frac{1}{2\pi RC} \qquad \text{...(3)}$$

and $$\frac{R_3}{R_4} = 2 \qquad \text{...(4)}$$

Thus, a ratio of R_3 to R_4 greater than 2 will provide sufficient loop gain for the circuit to oscillate at the frequency calculated using Eq. (3).

EXERCISE

1. Draw the circuit diagram of a Hartley oscillator. Derive the conditions for the maintenance and frequency of oscillations.

2. Explain the principle of Colpitt's oscillator. Find the condition of sustained oscillations.

3. Draw the circuit diagram and explain the working of a Wien bridge oscillator. Find the condition for maintenance of oscillations and expression for frequency of oscillations.

4. What is a phase shift oscillator? Draw the circuit of a phase shift oscillator and explain the working. Also give the expression for the frequency. (*Mahatma Gandhi*, 1989)

5. With relevant circuit diagram discuss the principle of operation of a phase shift oscillator using transistor. Obtain expression for its frequency and for the condition for the circuit to oscillate. (*Calicut*, 1989)

6. Why is crystal oscillator used in ratio transmitter?

7. In a Hartley oscillator the tank coil has two sections of inductances 80 mH and 20 mH. The capacitor has a capacitance of 500 pF. Neglecting the mutual inductance of the coil, find its frequency of oscillation.

Solution : $L_1 = 80 \times 10^{-3}$ H, $L_2 = 20 \times 10^{-3}$ H, $C = 500 \times 10^{-12}$ F.

$$f = \frac{1}{2\pi\sqrt{C(L_1 + L_2)}} = \frac{1}{2 \times 3.14\sqrt{500 \times 10^{-12} \times 100 \times 10^{-3}}}$$

$$= 2.24 \times 10^4 \text{ Hz.}$$

8. In a Colpitt's oscillator the inductance and capacitances used in the tuned circuit are 50 mH, and 100 pF and 400 pF. Calculate the frequency of oscillations.

Solution : $C_1 = 100 \times 10^{-12}$ F, $C_2 = 400 \times 10^{-12}$ F, $L = 50 \times 10^{-3}$ H.

$$f = \frac{1}{2\pi\sqrt{\dfrac{LC_1C_2}{C_1 + C_2}}}$$

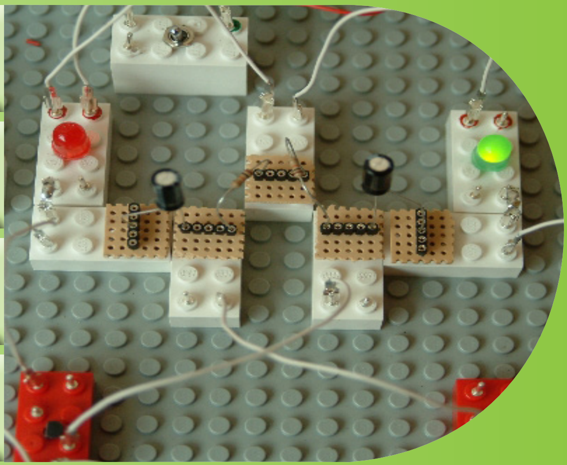
$$= \frac{1}{2 \times 3.14 \sqrt{\dfrac{50 \times 10^{-3} \times 100 \times 10^{-12} \times 400 \times 10^{-12}}{100 \times 10^{-12} + 400 \times 10^{-12}}}} = 8 \times 10^{4} \text{ Hz.}$$

9. The two capacitors in a Colpit's oscillator have values of 0.2 and 0.02 μF. The frequency of the oscillator is 10 kHz. Find the value of the inductor used in the circuit. (**Ans.** 13.93 mH)

10. The frequency of a Wien bridge oscillator is 3 kHz. If the value of the resistor in the bridge network is 200 kΩ, find the value of the capacitors.

Solution : $$f = \frac{1}{2\pi RC} \text{ or } 3 \times 10^{3} = \frac{1}{2\pi \times 200 \times 10^{3} C}$$

∴ Capacitance = C = 265.5 pF.



CHAPTER 67

MULTIVIBRATORS AND LOGIC GATES

AT A GLANCE			
67.1	Astable Multivibrator	**67.2**	Monostable Multivibrator
67.3	Bistable Multivibrator or Flip-Flop Circuit	**67.4**	The RS Flip-Flop
67.5	The J–K Flip-Flop	**67.6**	The NOR Gate
67.7	NOR Gate is a Universal Gate	**67.8**	The NAND Gate
67.9	NAND Gate is Universal Gate	**67.10**	The EXCLUSIVE OR Gate

67.1 Astable Multivibrator

An astable or free running multivibrator generates square waves of its own *i.e.,* without any external excitation. It has no stable state. It has only two quasistable states between which it keeps on oscillating of its own accord.

Circuit Details

Fig. 67.1 shows the circuit of an astable multivibrator. Q_1 and Q_2 are two identical transistors.

It is a two-stage *RC* coupled-amplifier. The output of each stage is coupled to the input of the other through a capacitor [C_1 or C_2]. Each stage will produce a phase shift of 180°. So the total phase shift between the input and output is 360°. Thus a positive feedback takes place and the circuit oscillates. The feedback is so strong that either the transistors are driven to saturation or to cut-off.

Astable Multivibrator.

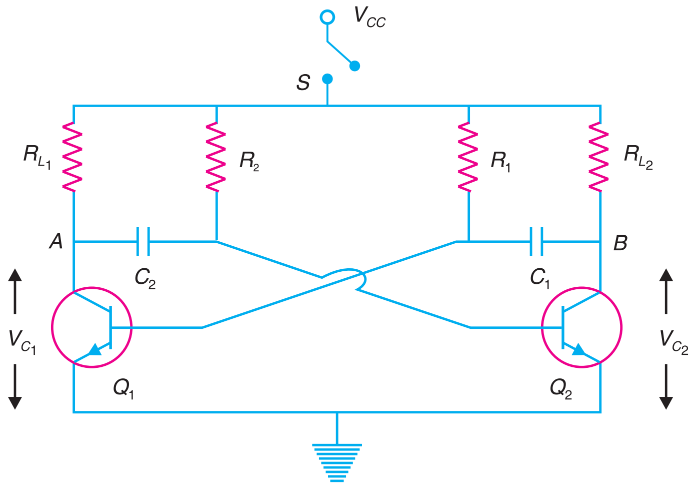


Fig. 67.1

Circuit operation

(*i*) When power V_{CC} is applied by closing switch *S*, collector current starts flowing in Q_1 and Q_2. The coupling capacitors C_1 and C_2 start charging up. The characteristics of no two similar transistors are identical. Suppose Q_1 conducts more than Q_2.

Then the collector current of Q_1 will rise rapidly. This will result in a decrease of the collector voltage of Q_1. The resulting negative signal is applied to the base of Q_2 through C_2. This drives Q_2 towards cut-off. Consequently, the collector voltage of Q_2 rises towards V_{CC}. This change in collector voltage of Q_2 (positive going signal) is fed to the base of transistor Q_1 through capacitor C_1. As a result of this positive going pulse, the collector current of Q_1 is further increased. The process being cumulative, in a short time, transistor Q_1 is saturated while Q_2 is cut-off. These actions are so rapid and instantaneous that C_1 does not get a chance to discharge. Under this situation, whole of V_{CC} drops across R_{L1} (since Q_1 is saturated or is on ON state) *i.e.,* $V_{C1} = 0$ and point *A* is at ground (or zero) potential. Also, since Q_2 is cut-off (OFF state), there is no drop across R_{L2} and point *B* is at V_{CC}.

(*ii*) Capacitor C_2 now begins to discharge which decreases the reverse bias on base of transistor Q_2. Ultimately a forward bias is re-stabilized at Q_2. Therefore, Q_2 begins to conduct. Consequently, collector of Q_2 becomes less positive. This negative going voltage signal is applied to the base of transistor Q_1 through the capacitor C_1. As a result, Q_1 is pulled out of saturation and is soon driven to cut-off. Simultaneously Q_2 is driven to saturation. Now V_{C2} decreases and becomes almost zero volt when Q_2 gets saturated. Consequently, potential of point *B* decreases from V_{CC} to almost zero volt.

The transistor Q_1 remains cut off and Q_2 in conduction until capacitor C_1 discharges through R_1, enough to decrease the reverse bias of Q_1.

The whole cycle repeats.

The output of the multivibrator is taken from the collector of either transistor. The output is a *square wave* (Fig. 67.2) with a peak amplitude equal to V_{CC}.

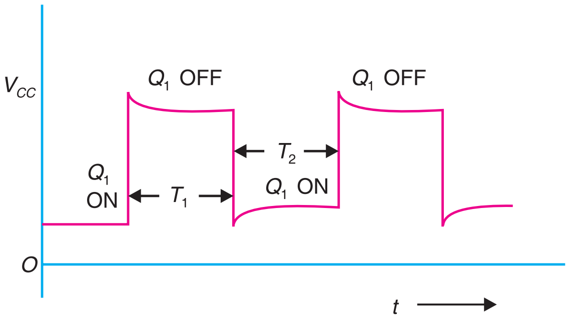


Fig. 67.2

Switching Times

The multivibrator circuit alternates between a state in which Q_1 is ON and Q_2 is OFF and a state in which Q_1 is OFF and Q_2 is ON.

ON time for Q_2 (or OFF time for Q_1) is

$$T_1 = 0.69\, R_1 C_1 \qquad \text{...(1)}$$

ON time for Q_1 (or OFF time for Q_2) is

$$T_2 = 0.69\, R_2 C_2 \qquad \text{...(2)}$$

Total period of the wave

$$T = T_1 + T_2 = 0.69\,(R_1 C_1 + R_2 C_2) \qquad \text{...(3)}$$

If $R_1 = R_2 = R$ and $C_1 = C_2 = C$ *i.e.,* the two stages are symmetrical,

$$T = 0.69\,(RC + RC) = 1.38\, RC \qquad \text{...(4)}$$

Frequency of oscillation

Frequency of oscillation, $f = \dfrac{1}{T} = \dfrac{1}{1.38RC} = \dfrac{0.7}{RC}.$

Here, f is in hertz, R in ohms and C in farads.

67.2 Monostable Multivibrator

It has *one stable state* and *one quasi-stable state. One transistor is always conducting* (*i.e.,* in the *ON* state) and *the other is non-conducting* (*i.e.,* in the OFF state). A triggering pulse is required to induce a transition from the stable state to the quasi-stable state. After certain time, the multivibrator automatically switches back from quasi-stable state to the stable state and remains in the stable state until another pulse is applied.

Circuit Diagram

Fig. 67.3 shows the circuit diagram of monostable multivibrator.

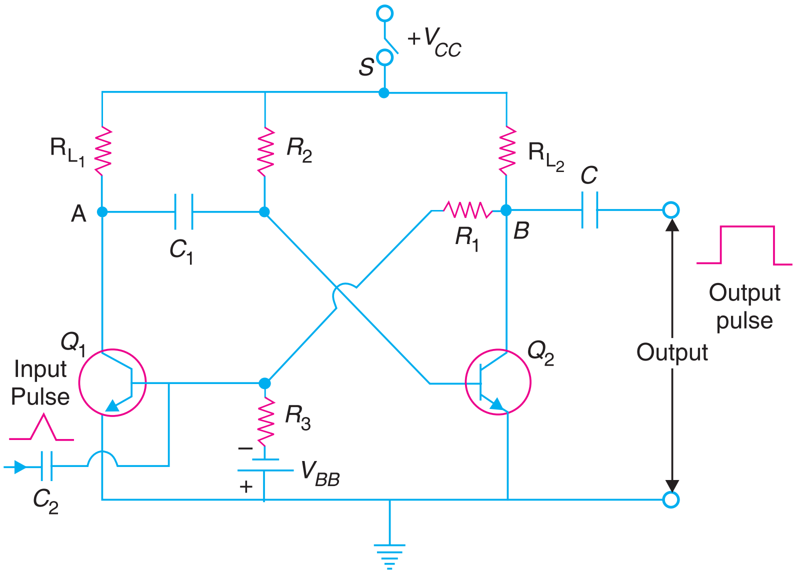
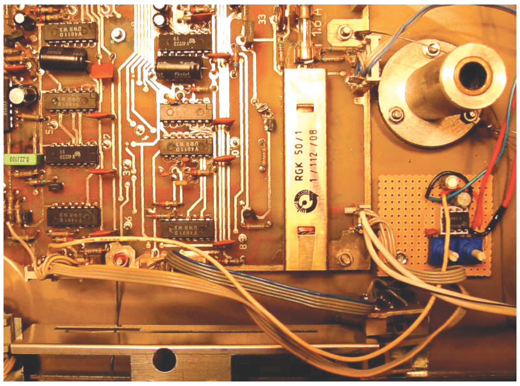


Fig. 67.3



Monostable Multivibrator.

Q_1 and Q_2 are two identical NPN transistors with equal collector loads *i.e.,* $R_{L1} = R_{L2}$. V_{BB} and R_3 reverse bias the transistor Q_1 and keep it cut-off. The supply voltage $+ V_{cc}$ and resistor R_2 forward bias the transistor Q_2 and keep it at saturation. The input triggering pulse is applied to the base of Q_1 through C_2. The output is taken from Q_2.

Working :

(*i*) Initially when switch S is closed and no triggering pulse is applied at C_2, Q_1 is cut-off and Q_2 conducts at saturation. This is the stable state of the multivibrator.

(*ii*) When a positive trigger pulse of short duration and of sufficient amplitude is applied to the base of Q_1 through C_2, it forward biases Q_1. Hence Q_1 starts conducting. Consequently, collector voltage of Q_2 falls due to voltage drop across R_{L1}. Thus potential of point A falls which is equivalent to a negative going voltage. This negative potential at the collector of Q_1 is fed to the base of Q_2 through capacitor C_1. Hence the forward bias of Q_2 decreases. Now the collector current of Q_2 decreases. Consequently, potential of point B increases due to lesser drop over R_{L2}.

The potential of B is applied to the base of Q_1 through R_1. So the forward bias of Q_1 increases. Now the transistor Q_1 conducts more and more and the potential of point A approaches zero volt. Quickly Q_1 is driven to saturation and Q_2 to cut-off. This is the quasi-stable state.

(*iii*) With Q_1 at saturation and Q_2 at cut-off, the circuit will return to initial stable state (*i.e.* Q_2 at saturation and Q_1 at cut-off) after some time.

During the time when Q_1 changes its state from 'OFF' to 'ON' the capacitor C_1 is charged fully. Since point A is almost at zero-volt, capacitor C_1 starts to discharge through saturated Q_1 to ground. It decreases the negative potential at the base of Q_2. As C_1 discharges further, transistor Q_2 is pulled out of cut-off. As Q_2 conducts further, a negative going voltage from point B via R_1 drives Q_1 into cut-off.

Thus the circuit returns to its initial stable state with Q_2 conducting at saturation and Q_1 cut-off.

To change this state, another pulse is to be applied.

The entire cycle repeats itself.

The width of the output pulse is determined by the time constant of C_1R_2 and given by

$$T = 0.69\ C_1R_2.$$

MMV is used to generate new and sharp pulses from distorted and used pulses in computers and tele-communication systems.

67.3 Bistable Multivibrator or Flip-Flop Circuit

It has *two absolutely stable states.* The circuit can stay in any one of the states indefinitely. When an external pulse is applied, the circuit changes from one stable state to another stable state. When another triggering pulse is applied only then it goes back to its original state.

The multivibrator is also known as *flip-flop.* The reason is that one triggering pulse causes the multivibrator to 'flip' from one stable state to another stable state while the second pulse causes it to 'flop' back to original stable state.

Circuit Diagram

Fig. 67.4 shows the circuit diagram of bistable multivibrator.

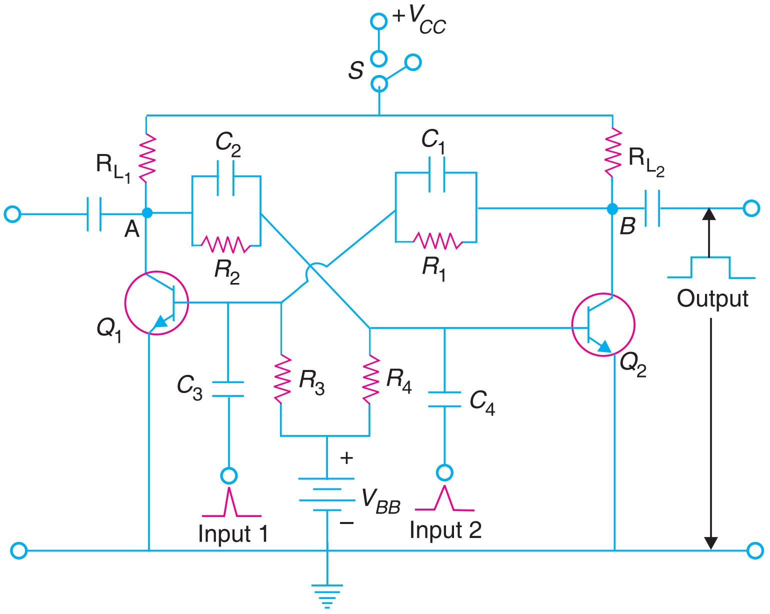
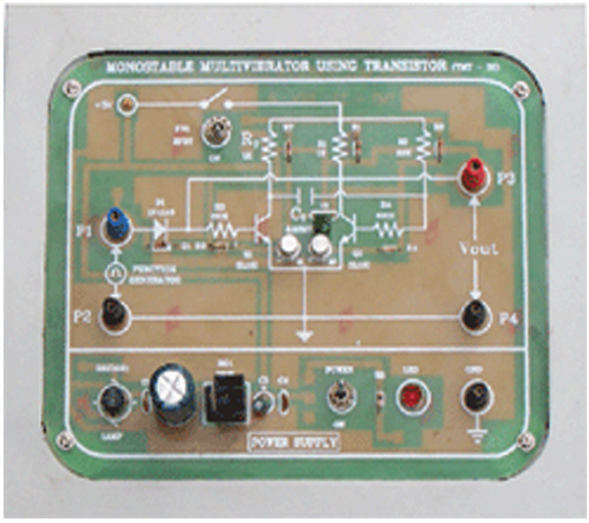


Fig. 67.4

It consists of two identical transistors Q_1 and Q_2 in common-emitter configuration, in such a way that the output of one is given as the input of the other. The feedback is given through the resistors R_1, R_2 shunted by capacitors C_1, C_2. The main purpose of the capacitors C_1 and C_2 is to improve the switching speed of the circuit. The bias battery V_{BB} with resistors R_3, R_4 forward biases the bases of Q_1 and Q_2. Suitable triggering pulses are applied to the bases of Q_1 and Q_2 through C_3 and C_4. The output can be taken across either Q_1 or Q_2.



Bistable Multivibrator.

Circuit Operation

(i) When the supply voltage V_{CC} is applied by closing switch S, collector current starts flowing in transistors Q_1 and Q_2. Since no two transistors can have the same characteristics, one transistor, say Q_1, will conduct more rapidly than the other. Then the collector current of Q_1 will rise at a faster rate causing a decrease in its collector voltage. The resulting negative signal is applied to the base of Q_2 through C_2 and drives it towards cut-off. Consequently, the collector voltage of Q_2 rises towards V_{CC}. This positive going signal is fed to the base of Q_1 through C_1. As a result, the collector current of Q_1 is further increased. The process being cumulative, in a short time, transistor Q_1 is saturated (ON) while Q_2 is cut-off (OFF). This is the *first stable state* of the multivibrator. When the circuit is left free, it will remain in this stable state indefinitely.

(*ii*) To change the first stable state into second stable state, a negative trigger pulse is applied to the base of Q_1 through C_3. This reduces the forward bias of transistor Q_1. As a result, the collector current decreases or the collector voltage increases.

This rising collector voltage is fed to the base of Q_2 where it forward biases the base-emitter junction of Q_2, increases its collector current and decreases collector voltage. This decreasing collector voltage of Q_2 is fed to the base of Q_1 where it reverse biases base-emitter junction of Q_1 and decreases its collector current.

After few cycles, Q_2 is driven to saturation (ON) and Q_1 to cut off (OFF).

This is the *second stable state* of the multivibrator.

When the circuit is left free, it will remain in this stable state indefinitely, until another pulse is applied.

The output waveforms are shown in Fig. 67.5.

Uses (*i*) It is used as "Flip-Flop" in digital electronics to store one 'bit'. The multivibrator is used in many digital operations such as counting and storing the binary information.

(*ii*) It is used in pulse generation.

(*iii*) It is used as a frequency divider in timing circuits.

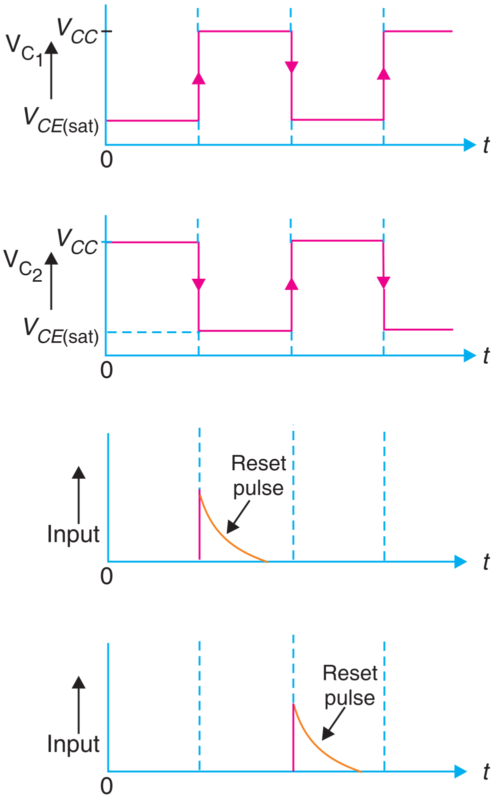


Fig. 67.5

67.4 The RS Flip-Flop

A bistable multivibrator or flip-flop has two stable states and can remain in either of these two states unless an external trigger pulse switches it from one state to the other. Thus it can stay in one of the two possible states after an input has been applied. The state does not change following the removal of the input. Therefore, a flip-flop is essentially a 1-bit *memory* or *storage device*.

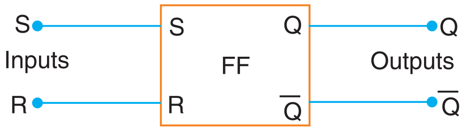


Fig. 67.6

Fig. 67.6, shows the block diagram of the *R–S* flip-flop. It has two inputs: the *S* (Set) and *R* (Reset). The two outputs denoted by Q and $\overline{Q}$ are complementary to each other. Whenever, Q is at the state 1, the $\overline{Q}$ is at the state O and vice versa. The value of Q is the state of the flip-flop. Both input and output signals consist of voltage levels which correspond to 0s and 1s.

Implementation of *R–S* Flip Flip using two *NOR* Gates

The *R–S* flip-flop can be implemented using two *NOR* gates as shown in Fig. 67.7.

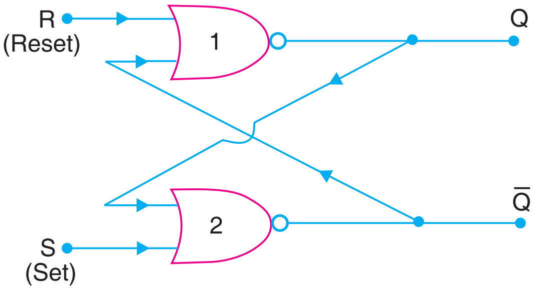


Fig. 67.7

The *NOR* gates (under positive logic) are inter-connected so that the output of *NOR* – 1 is fed as one of the inputs to *NOR*–2 and vice versa.

(*i*) When the input is switched as $R = 1$ and $S = 0$, the output of *NOR* – 1 is 0 (at Q). Hence both the inputs to the *NOR* –2 are zero. This makes the output $\overline{Q}$ as 1. Thus when $R = 1$ and $S = 0$, we have $Q = 0$. The state of the *FF* is 'Reset' and remains in the stable state with $Q = 0$.

(*ii*) When the input is switched to $R = 0$ and $S = 1$, the output of *NOR* –2 is forced to be 0. Hence both the inputs to the *NOR* – 1 are zero, making the output Q as 1. Thus, when $R = 0$ and $S = 1$, we have $Q = 1$ and the *FF* is 'Set' and remain in the stable state with $Q = 1$.

(*iii*) When both input lines R and S carry 0 signal, the flip-flop remains in the same state. Thus when $R = 0$ and $S = 0$, Q = unchanged.

(*iv*) The input condition $R = 1$ and $S = 1$ is not allowed because then the outputs Q and $\overline{Q}$ both will become zero–a condition which is contradictory to the assumption that Q and Q are complementary.

The above output/input correspondence can be represented by a Truth Table as given below:

TRUTH TABLE

Inputs	*Outputs*			*Comment*
R	S	Q_{n+1}	$\overline{Q}_{n+1}$	
0	0	Q_n	$\overline{Q}_n$	No change
1	0	0	1	Reset
0	1	1	0	Set
1	1	Not Allowed		

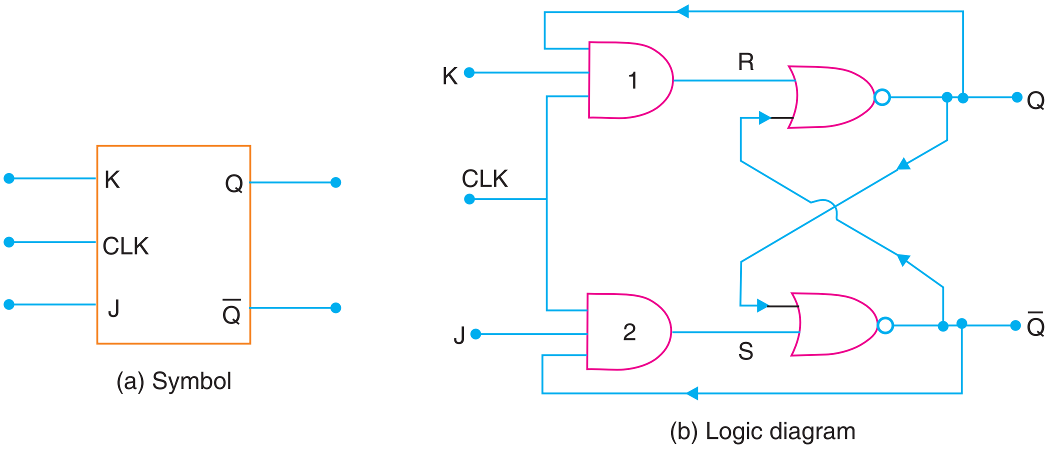
67.5 The J–K Flip-Flop

A *JK* flip-flop is a refinement of the *RS* flip-flop in that the indeterminate state (Inputs $R = 1$, $S = 1$) of the *RS* type is defined in the *JK* type. Inputs J and K behave like inputs S and R to set and clear the flip-flop, respectively. The input marked J is for *set* and the input marked K is for *reset*.When both inputs J and K are equal to 1, the flip-flop switches to its complement state, that is, if $Q = 1$, it switches to $Q = 0$, and vice versa.

J–K Flip-Flop.

Construction

Fig. 67.8 shows the *JK* flip-flop constructed with two cross-coupled *NOR* gates and two AND gates.



(a) Symbol

(b) Logic diagram

Fig. 67.8

Each *AND* gate has three input terminals. Both the *AND* gates are driven by clock pulses. J and K are the control inputs (data inputs). Q and $\overline{Q}$ are the outputs.

Operation

When the clock pulses are absent, the two *AND* gates are disabled and so no change of state in the output takes place. What happens when the clock pulses are present ?

(i) Inactive: Assume that $Q = 1$, and $\overline{Q} = 0$ initially. Now, let the inputs be $J = 0$, $K = 0$. Both input gates are disabled and the circuit is inactive at all times including the rising edge of the clock. Therefore, no change takes place in the output. i.e., Q remains 1 and $\overline{Q}$ remains 0 (initial state).

(ii) Reset: When the inputs are $J = 0$, $K = 1$, on arrival of a clock pulse, the *AND* – 1 is enabled and the *AND* – 2 is disabled (since $J = 0$). Since the output of *AND* – 2 is 0, the flip-flop is not set.

With Q initially high and $K = 1$, the *AND* – 1 passes a reset trigger as soon as the positive clock edge arrives. This forces $Q = 1$ and $Q = 0$. Therefore, $J = 0$ and $K = 1$ means that the rising clock edge resets the *FF*. i.e., the state of the *JK* flip-flop changes to 0. (If the initial state is already $Q = 0$, then it will continue to remain so, even in the presence of the clock pulse).

(iii) Set: When the inputs are $J = 1$ and $K = 0$, with $Q = 0$, *AND* gate – 1 is disabled and *AND* – 2 is enabled and the *FF* passes a set-trigger during the positive clock edge. This drives Q into the 1 state i.e., $Q = 1$ and $\overline{Q} = 0$ i.e., $J = 1$, $K = 0$ means that the next positive clock edge sets the *FF*. i.e., the state of the *FF* changes to 1.

(iv) Toggle : When the inputs are $J = 1$ and $K = 1$, it is possible to set or reset the flip-flop, depending on the existing state of the outputs.

Suppose Q is 1 (existing state) and *AND*-1 passes a reset trigger on the next positive clock edge. This makes $Q = 0$ and $\overline{Q} = 1$.

When the existing state $Q = 0$, $\overline{Q} = 1$, the *AND*-2 passes a set trigger on arrival of the next positive clock edge. This makes $Q = 1$ and $\overline{Q} = 0$.

In either case, the output state of the flip-flop is complemented.

Thus a change of state has occurred. This state will change each time a clock pulse arrives if $J = K = 1$. It means that *flip-flop toggles from one state to the other at each clock pulse.*

The *Truth Table of JK* flip flop is given below. It gives the output with the application of clock pulse as a function of the data inputs J and K.

TRUTH TABLE

Inputs			*Output*	*Comment*
Clock	*J*	*K*	Q_{n+1}	
⎍	0	0	Q_n	No Change
”	0	1	0	Reset
”	1	0	1	Set
”	1	1	toggle	To opposite state

67.6 The NOR Gate

A negation following an OR is called a NOT – OR or a NOR gate.

It is the circuit combination in which a NOT circuit follows an OR gate.

A NOR gate can be implemented by placing a transistor ‘NOT’ circuit after the diode ‘OR’ circuit (Fig. 67.9).

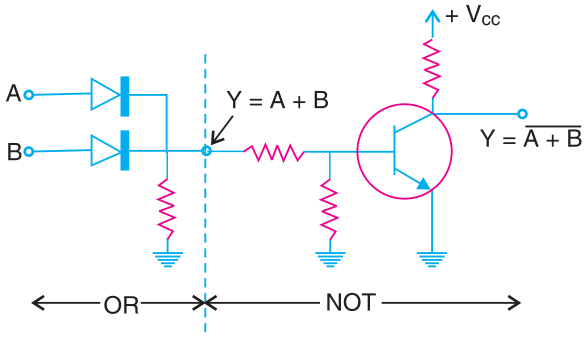


Fig. 67.9

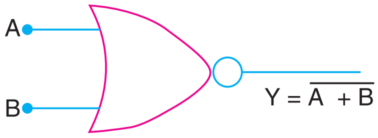


Fig. 67.10

The symbol of a NOR gate is shown in Fig 67.10.

The Boolean expression for the NOR gate is $Y = \overline{A + B}$.

The output of the OR gate is inverted by the NOT gate and thus yields NOR operation.

Truth Table. Consider the input values as follows:

Case 1 : When $A = 0$ and $B = 0$, then $A + B = 0$.

$\therefore \quad Y = \overline{A + B} = 1.$

Case 2 : When $A = 0$ and $B = 1$, then $A + B = 1$.

$\therefore \quad Y = \overline{A + B} = 0.$

Case 3 : When $A = 1$ and $B = 0$, then $A + B = 1$.

$\therefore \quad Y = \overline{A + B} = 0.$

Case 4 : When $A = 1$ and $B = 1$, then $A + B = 1$.

$\therefore \quad Y = \overline{A + B} = 0.$

Truth Table – NOR Operation

Input		*Output*
A	B	$Y = \overline{A + B}$
0	0	1
0	1	0
1	0	0
1	1	0

67.7 NOR Gate is a Universal Gate

A NOR gate can be used to realize the basic logic functions: OR, AND and NOT. So NOR gate is also called **universal gate**. Thus, in digital circuits, NOR gate serves as a building block.

If all the three basic gates can be constructed using NOR gates alone, then NOR can be called a universal gate.

1. As OR gate

An 'OR gates can be constructed using 'NOR' gates as follows.

Consider a two input 'NOR' gate (N_1) and a single input 'NOR' gate N_2 (Fig. 67.11).

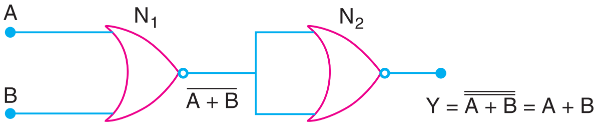


Fig. 67.11

By connecting the output of two input 'NOR' gate (N_1) as the input of the single input 'NOR' gate (N_2), an 'OR' gate is constructed. The output of N_1 is $\overline{A + B}$, and this is used as the input of N_2. Hence the output of N_2 is

$$Y = \overline{\overline{A + B}} = A + B \left(\because \overline{\overline{A}} = A\right).$$

Thus the combination of the 'NOR' gates N_1 and N_2 as shown in Fig. 67.11 acts as an 'OR'gate.

2. As AND gate

An 'AND' gate can be constructed using 'NOR' gate as follows.

Consider two single input 'NOR' gates N_1 & N_2 and a two input 'NOR' gate N_3. The outputs of the NOR gates N_1 & N_2 are used as the inputs of N_3 (Fig. 67.12).

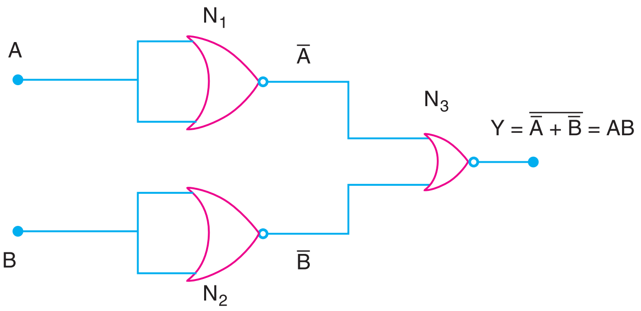


Fig. 67.12

The outputs of N_1 & N_2 are $\overline{A}$ and $\overline{B}$. These are used as the inputs of N_3. Hence the output of N_3 is

$$Y = \overline{\overline{A} + \overline{B}} = \overline{\overline{A}} \, . \, \overline{\overline{B}} \text{ (By De Morgan's Theorem)}$$

$$= A.B \left(\because \overline{\overline{A}} = A\right)$$

Thus the combination of the 'NOR' gates N_1, N_2 and N_3 as shown in Fig. 67.12, acts as an 'AND' gate.

3. As NOT gate

(*i*) The two inputs of NOR gate are tied together, as shown in Fig. 67.13(a).

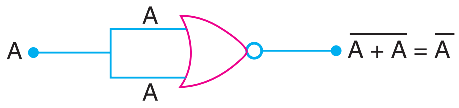


Fig. 67.13 (a)

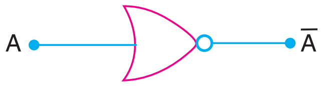


Fig. 67.13 (b)

The output is $\overline{A + A}$

By De Morgan's Theorem,

$$\overline{A + A} = \overline{A}.$$

(*ii*) When only single input is used, the circuit symbol of Fig. 67.13 (b) is widely used.

67.8 The NAND Gate

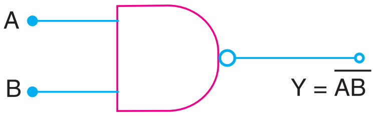


Fig. 67.14

A negated AND is called a NOT-AND or NAND gate. It is obtained by connecting a NOT gate in the output of an AND gate.

Fig. 67.14 gives the logic symbol for two input NAND gate.

The output of the AND gate is inverted by the NOT gate.

Its output is given by the Boolean equation $Y = \overline{AB}$.

The NAND gate is implemented by placing a transistor NOT circuit after the diode AND circuit (Fig. 67.15).

Truth Table. Consider the input values as follows :

Case 1: When $A = 0$ and $B = 0$, then $AB = 0$.

$\therefore \quad Y = \overline{AB} = 1.$

Case 2 : When $A = 0$ and $B = 1$, then $AB = 0$.

$\therefore \quad Y = \overline{AB} = 1.$

Case 3 : When $A = 1$ and $B = 0$, then $AB = 0$.

$\therefore \quad Y = \overline{AB} = 1.$

Case 4 : When $A = 1$ and $B = 1$, then $AB = 1$.

$\therefore \quad Y = \overline{AB} = 0.$

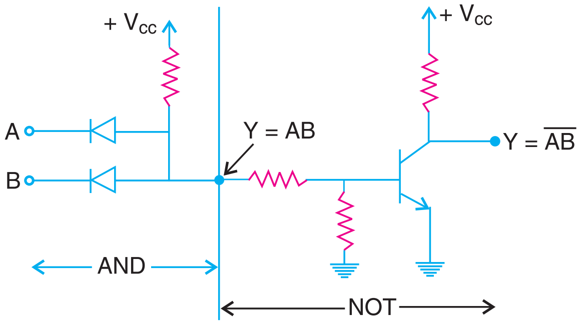


Fig. 67.15

Truth Table – NAND Operation

Input		*Output*
A	B	$Y = \overline{A \bullet B}$
0	0	1
0	1	1
1	0	1
1	1	0

Thus this gate *gives output* 1 *if either A or B or both are* 0.

67.9 NAND Gate is a Universal Gate

A NAND gate is known as a universal gate because it can be used to realize all the three basic logic functions of an OR gate, AND gate and NOT gate. Thus, in digital circuits, NAND gate serves as a building block.

1. As OR Gate

OR gate can be made out of the three NAND gates (Fig. 67.16).

However, the OR function may not be very clear from the figure because De Morgan's theorem is needed to prove that $\overline{\overline{A} \bullet \overline{B}} = A + B$.

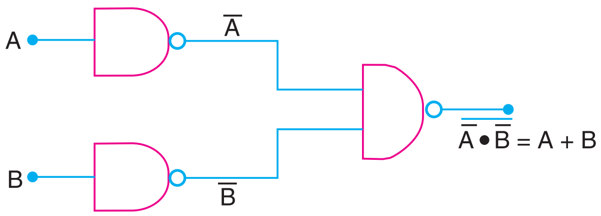

A
$\overline{A}$
B
$\overline{B}$
$\overline{\overline{A} \bullet \overline{B}} = A + B$

Fig. 67.16

2. As AND Gate

An AND gate can be produced by using two NAND gates (Fig 67.17).

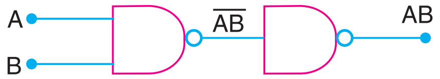

A
B
$\overline{AB}$
AB

Fig. 67.17

3. As NOT Gate

(*i*) If the two inputs of a NAND gate are connected together, then we get a NOT gate (Fig. 67.18a).

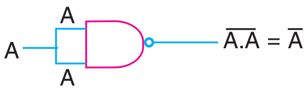

A
A
A
$\overline{A.A} = \overline{A}$

Fig. 67.18(a)

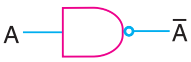

A
$\overline{A}$

Fig.67.18(b)

(*ii*) When only single input is used, the circuit symbol of Fig. 67.18b is used.

67.10 The EXCLUSIVE OR Gate

The EXCLUSIVE OR (abbreviated as XOR) operation obeys the definition that a two input circuit provides an output when one input or the other is present but not when both inputs are present.

The XOR operation may also be stated as, "If $A = 1$ and $B = 0$, or if $B = 1$ and $A = 0$, then $Y = 1$".

The Boolean expression is

$$Y = A\overline{B} + B\overline{A} .$$

The exclusive – OR operation is denoted by ⊕.

$$Y = A \oplus B = A\overline{B} + B\overline{A} .$$

The block diagram according to this logic is given in Fig. 67.19.

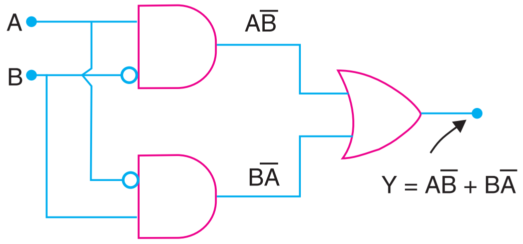

A
B
$A\overline{B}$
$B\overline{A}$
$Y = A\overline{B} + B\overline{A}$

Fig. 67.19

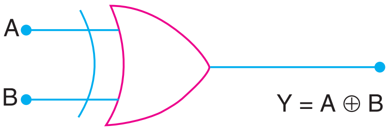

A
B
$Y = A \oplus B$

Fig. 67.20

It is constructed using two NOT gates, two AND gates and one OR gate.

The logic symbol for the XOR gate is shown in Fig. 67.20.

Truth table. We take the input values as follows:

Case 1 : When A = 0, B = 0; so that $\overline{A}=1$ and $\overline{B}=1$.

Then $Y = 0 \cdot 1 + 1 \cdot 0 = 0 + 0 = 0.$

Case 2 : When $A = 0, B = 1$; so that $\overline{A}=1$ and $\overline{B}=0$.

Then $Y = 0 \cdot 0 + 1 \cdot 1 = 0 + 1 = 1.$

Case 3 : When $A = 1, B = 0$; so that $\overline{A}=0$ and $\overline{B}=1$.

Then $Y = 1 \cdot 1 + 0 \cdot 0 = 1 + 0 = 1.$

Case 4 : When $A = 1, B = 1$; so that $\overline{A}=0$ and $\overline{B}=0$.

Then $Y = 1 \cdot 0 + 0 \cdot 1 = 0 + 0 = 0.$

Truth Table for XOR Gate

Input		***Output***
A	B	$Y = A\overline{B} + B\overline{A}$
0	0	0
0	1	1
1	0	1
1	1	0

Thus output is 1 only when one of the two inputs is 1 but not both. Thus it *produces an output when the two inputs are different.*

EXERCISE

1. Draw a circuit of astable multivibrator and explain its operation. What is the frequency of the circuit? (*Madurai,* 87)
2. In an astable multivibrator, the value of $R_1 = R_2 = 15\ k\Omega$, and $C_1 = C_2 = 0.005\ \mu F$. Calculate the frequency of oscillation. (*Andhra*, 1995)

Sol. $$f = \frac{1}{1.38RC} = \frac{1}{1.38 \times (15 \times 10^3) \times (0.005 \times 10^{-6})} = 9.66 \times 10^3 \text{ Hz.}$$

CHAPTER

68

P9NK60ZD
B9NK60ZD
F9NK60ZD

METAL OXIDE SEMICONDUCTOR FIELD EFFECT TRANSISTOR (MOSFET)

AT A GLANCE

68.1	Introduction
68.2	The Depletion MOSFET
68.3	Static Characteristics of Depletion MOSFET
68.4	The Enhancement MOSFET
68.5	Characteristics of Enhancement MOSFET
68.6	MOSFET Biasing

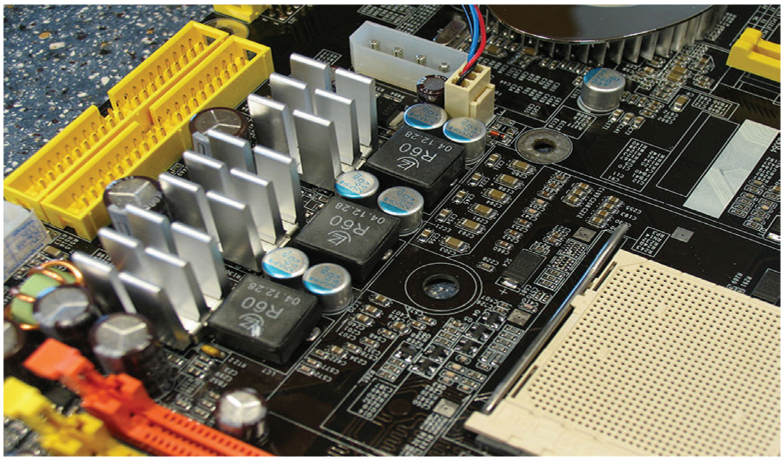
68.1 Introduction

There are two types of MOSFETs.

(*i*) The Depletion MOSFET.

(*ii*) The Enhancement MOSFET.

In both types, *p*-channel and *n*-channel versions exist.



68.2 The Depletion MOSFET

Construction

Fig. 68.1 shows the construction of an *n*-channel depletion MOSFET. It consists of a lightly doped *p*-type substrate into which two heavily-doped n^+-regions are diffused. These two n^+ regions act as source *S* and drain *D*.

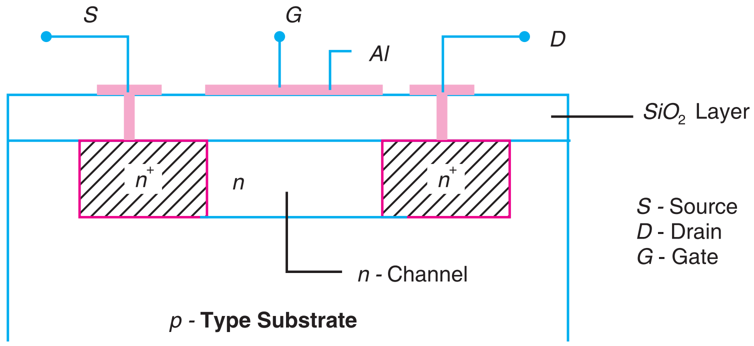


Fig. 68.1

They are separated by about 5 μm. A lightly-doped *n*-channel is diffused between the source and the drain. A thin (~ 0.1 μm) layer of insulating silicon dioxide (SiO_2) is grown over the surface of the structure. Holes are cut into the oxide layer to make metallic contacts with the source and the drain. Then a metallic layer (say, of aluminium) is overlaid on the oxide layer, covering the entire channel region. This aluminium layer acts as gate (*G*). Simultaneously, aluminium contacts are made to the source and the drain.

Fig. 68.2 shows the symbol of *n*-channel depletion-type MOSFET. Usually, the substrate terminal *SS* is internally connected to the source terminal *S*.

Operation

Suppose a potential is applied to the gate such that it is positive with respect to source. Then, the positive charge on the gate attracts additional free electrons into the channel from the source. This enhancement of mobile carriers decreases the channel resistance.

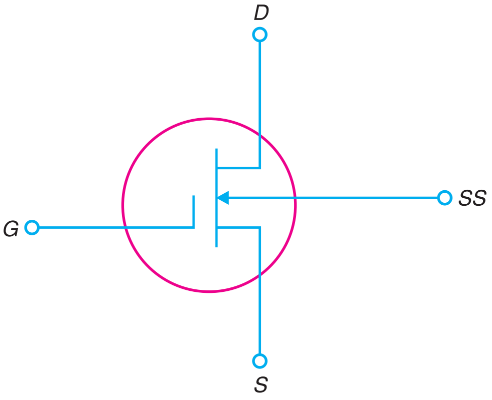


Fig. 68.2

SI
DN2540N5
9511
THAILAND
T93XB
500Ω 10%
3 0110

Depletion MOSFET.

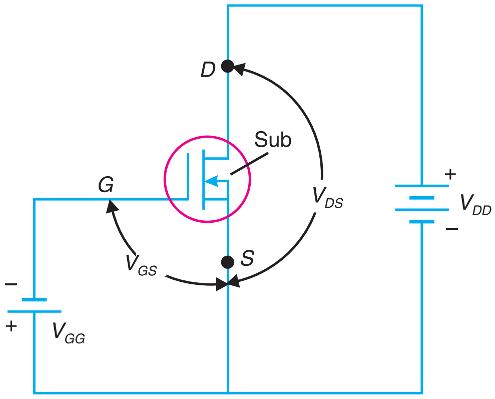


Fig. 68.3

Suppose the applied voltage makes the gate negative with respect to source (Fig. 68.3). Then, the negative charge on the gate forces free electrons out of the channel.

A carrier-depletion region is formed on the surface of the silicon at the oxide-silicon interface. The channel is thus constricted by a negative gate voltage. The resistance of the channel increases. When the gate is sufficiently negative, the depletion region extends completely across the channel and joins with depletion region of the *p-n* junction on the other side of the channel. Under this condition, the channel cannot conduct current between drain and source. Thus drain and source become cut off. The negative gate voltage when channel becomes nonconducting is the pinch off voltage V_P.

A positive voltage V_{DS} is also applied across the drain *D* and the source *S* (Fig. 68.3). There is a voltage drop along the channel, with the drain end of the channel positive relative to the source end. This further increases the depletion region at the surface.

68.3 Static Characteristics of Depletion MOSFET

1. Output or Drain Characteristics. Fig. 68.4 shows the drain characteristic curves (i_D versus V_{DS} at constant V_{GS}) for an *n*-channel MOSEFT.

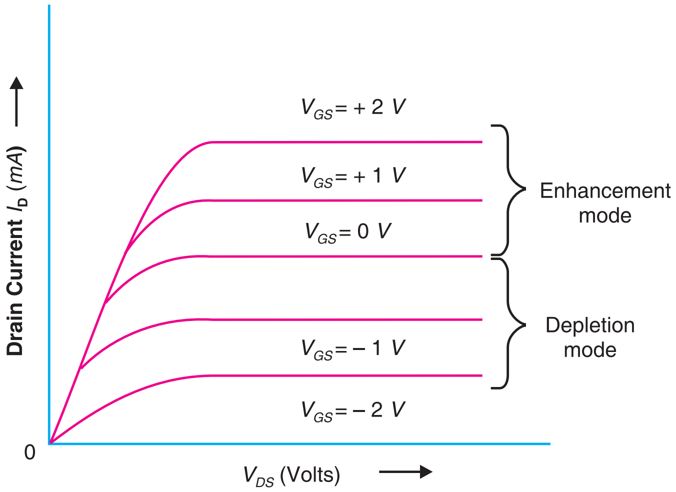


Fig. 68.4

An *n*-channel MOSFET may be operated in either the enhancement mode or the depletion mode. The enhancement mode occurs for positive values of V_{GS} while the depletion mode occurs for negative values of V_{GS}.

When no gate-to-source voltage is applied ($V_{GS} = 0$), a significant drain current i_D flows due to the flow of majority carriers (electrons) in the *n*-channel from source to drain under V_{DS}. As V_{DS} increases, i_D increases to a saturation value.

When the gate is made negative, positive charges are induced in the channel through the dielectric SiO_2. This causes depletion of electrons (majority carriers) in the channel. Therefore, the channel becomes less conductive. Consequently, the drain current progressively decreases to almost zero as V_{GS} is made more and more negative, at all values of V_{DS}.

When the gate is made positive, negative charges are induced in the channel, thus enhancing the majority-carriers. Hence the drain current increases above its value at $V_{GS} = 0$, for all V_{DS}.

Thus, this MOSFET can be operated in depletion as well as in enhancement mode although it is termed depletion type.

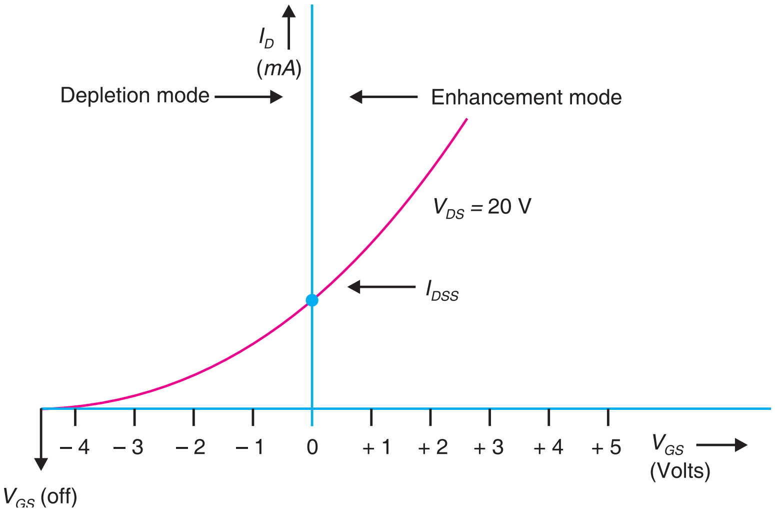


Fig. 68.5

2. Transfer Characteristics. The variation of drain current I_D with gate voltage V_{GS} at a constant drain voltage V_{DS} is represented by the 'transfer curve' for the MOSFET (Fig. 68.5). It may be seen that drain current I_D flows even when gate-bias $V_{GS} = 0$. When V_{GS} is made more and more negative, I_D goes on decreasing.

68.4 The Enhancement MOSFET

Construction

Fig. 68.6 shows the construction of an *n*-channel enhancement type MOSFET. It consists of a lightly doped substrate of *p*-type silicon. Two highly doped n^+ regions are diffused in the substrate. These two n^+ regions act as source (*S*) and drain (*D*). A thin layer of silicon dioxide (SiO_2) is deposited over the substrate. Then a thin film of metal aluminium is deposited over SiO_2. This aluminium layer acts as gate (*G*).

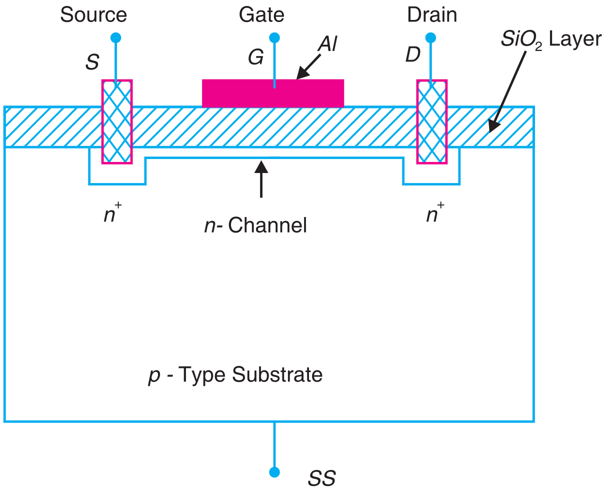
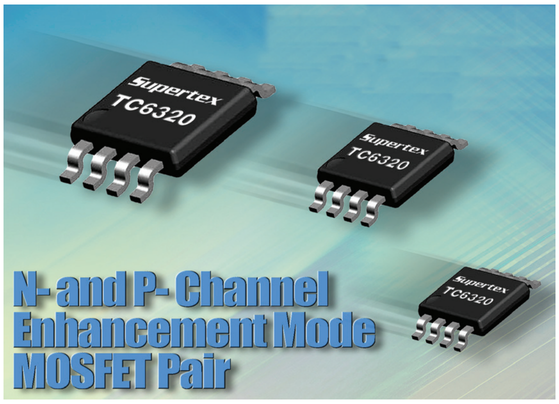


Fig. 68.6

Fig. 68.7 shows the symbol of *n*-channel E-MOSFET. The vertical line (representing channel) is broken to represent that there is no continuous channel in an enhancement MOSFET.



Enhancement MOSFET.

Operation

The metallic layer of *Al* and the upper surface of the substrate act as the parallel plates of a capacitor. The insulating layer of SiO_2 acts as the dielectric medium. When a *positive* voltage is applied to the gate *G* relative to source *S*, the capacitor begins to charge (Fig. 68.8). Consequently negative charges appear in the substrate between drain and source. Thus in effect an *n*-channel is created which allows the current to flow in source-to-drain circuit. Thus the MOSFET conducts only when gate is made positive relative to the source.

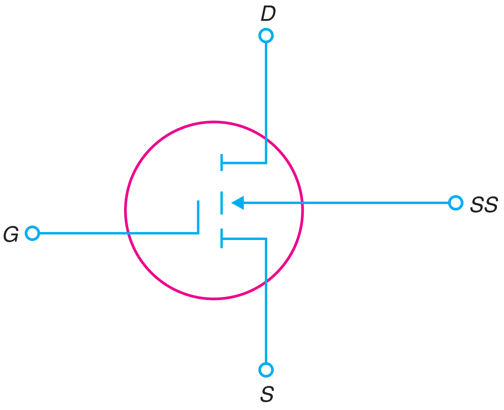


Fig. 68.7

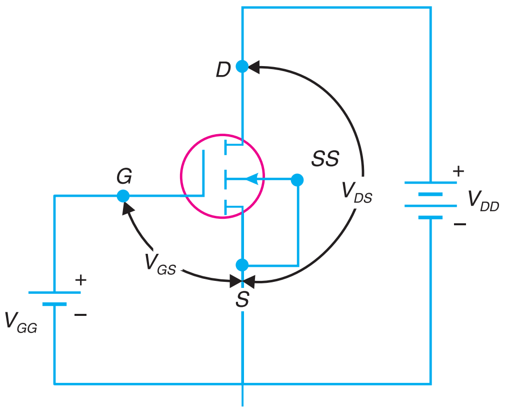


Fig. 68.8

The MOSFET can never operate with a negative gate voltage. The MOSFET is cut-off when $V_{GS} = 0$.

68.5 Characteristics of Enhancement MOSFET

1. Output or Drain Characteristics. Fig. 68.9 shows the drain characteristics of an *n*-channel enhancement MOSFET. Each curve shows the variation of drain current I_D with the drain-to-source voltage (V_{DS}) for a fixed value of gate-to-source voltage (V_{GS}).

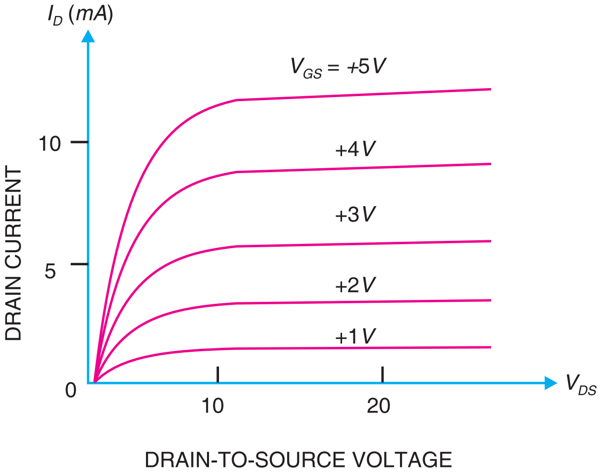


Fig. 68.9

The magnitude of the drain current I_D is a function of V_{DS} and V_{GS}. When $V_{GS} = 0$, $I_D = 0$ because two back to back *p–n* junctions between *S* and *D* (one between source *S* and the substrate and the other between the substrate and the drain *D*) are reverse biased regardless of the value of V_{DS}.

For a fixed positive V_{GS}, the drain current first increases rapidly with V_{DS} and then saturates. It increases for higher values of V_{GS}.

The minimum positive value of V_{GS} at which the drain current is established is called the 'gate-source threshold voltage' V_{GST} or V_T.

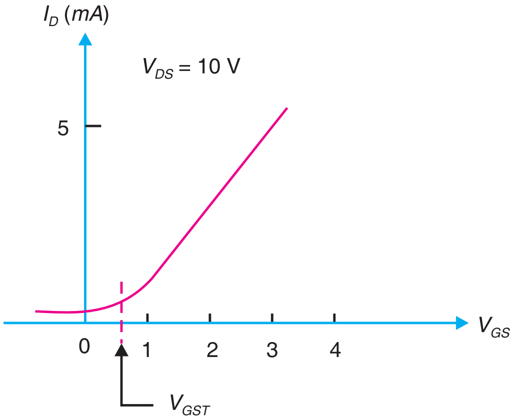


Fig. 68.10

2. Transfer Characteristics. The curve showing the variation of drain current I_D with gate-to-source voltage V_{GS} for a fixed value of V_{DS} is called the transfer characteristic. The transfer curve for an *n*-channel enhancement MOSFET is shown in Fig. 68.10. I_D flows only when V_{GS} exceeds gate-to-source threshold voltage V_T. With increase in V_{GS}, I_D increases slowly at first and then rapidly.

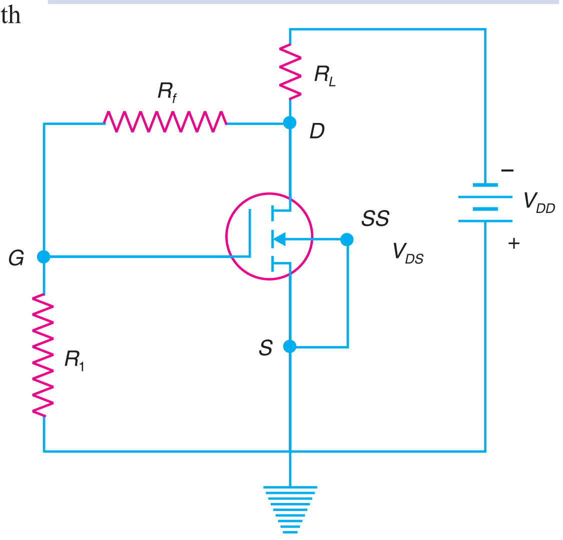


Fig. 68.11

68.6 MOSFET Biasing

An enhancement MOSFET requires forward biasing of the gate-to-source junction. Fig. 68.11 shows a circuit arrangement for biasing an enhancement MOSFET. It provides *forward biasing of the gate-to-source junction* whose magnitude is given by

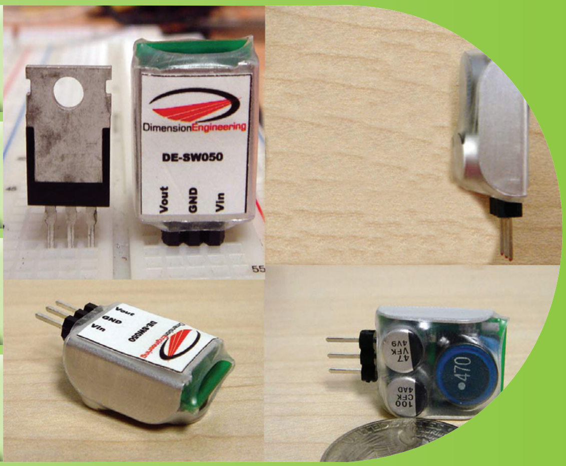
$$V_{GS} = \frac{R_1}{R_1 + R_f} V_{DS} .$$

Here the feedback resistor R_f provides the operating point stability.

EXERCISE

1. Sketch neatly the construction of a MOSFET using a P-type silicon bar. Explain how a channel may be developed in such a device. Draw and explain the typical drain current-drain voltage characteristics as a function of the gate voltage. (*Calcutta,* 81)
2. Sketch and explain the cross-section of an N-channel enhancement and depletion MOSFET. Draw the static drain characteristic and transfer characteristic curves for an N-channel MOSFET which may be operated in either the depletion or enhancement mode. (*Rohilkhand,* 88)
3. Explain the construction and working of Enhancement MOSFET, with a neat diagram. Draw its drain and transfer characteristics.
4. Draw a biasing circuit for an enhancement MOSFET.

CHAPTER 69



IC VOLTAGE REGULATORS

AT A GLANCE

69.1	Introduction
69.2	Fixed Voltage Regulators
69.3	Regulated Dual Supplies
69.4	Adjustable Voltage Regulator

69.1 Introduction

A voltage regulator is a circuit that gives a constant output voltage irrespective of the variations in input voltage and load current. IC regulators have three pins; one for the unregulated input voltage, one for the regulated output voltage, and one for ground. Some important types of linear *IC* voltage regulators are:

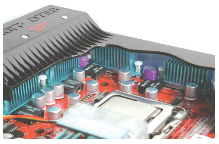
1. Fixed positive/negative output voltage regulators
2. Adjustable output voltage regulators.

69.2 Fixed Voltage Regulators

(*i*) The LM78XX series (where XX = 05, 06, 08, 10, 12, 15, 18 or 24) are three terminal, positive fixed voltage regulators. In 78XX, the last two numbers (XX) indicate the output voltage.

The 7805 produces an output of + 5 V, the 7806 produces + 6 V, the 7808 produces + 8 V, and so on, up to the 7824, which produces an output of + 24 V.

Fig. 69.1 (*a*) shows an LM7805 connected as a fixed voltage regulator. Pin 1 is the input, pin 2 is the output, and pin 3 is ground. The LM7805 has an output voltage of + 5 V and a maximum load current over 1A.



Fixed Voltage Regulator.

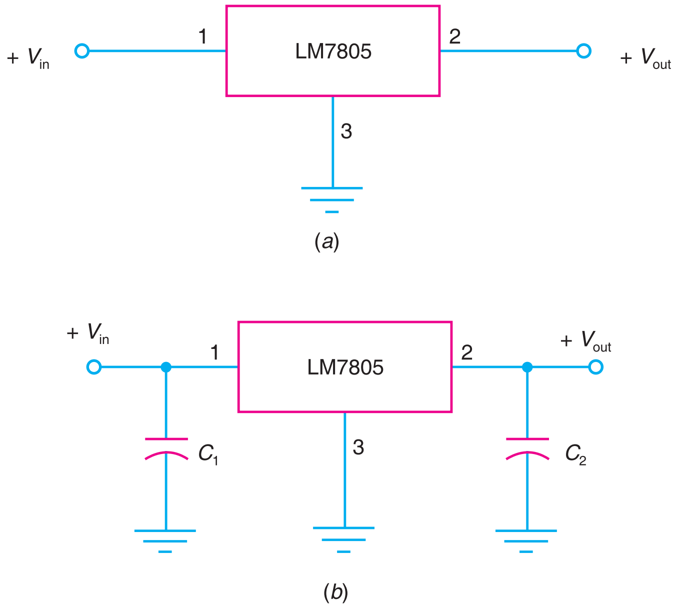

+ V_{in}
1
LM7805
2
+ V_{out}
3
(a)
+ V_{in}
1
LM7805
2
+ V_{out}
C_1
3
C_2
(b)

Fig. 69.1

The inductance of the connecting wire may produce oscillations inside the *IC*. The input capacitor C_1 on pin 1 (Fig. 69.1 b) prevents oscillations. To improve the transient response of the regulated output voltage, a bypass capacitor C_2 is used on pin 2.

(*ii*) The LM79XX series is a group of negative voltage regulators with preset voltages of – 5, – 6, – 8, – 10, – 12, – 15, – 18, or – 24 V. For instance, an LM7905 produces a regulated output voltage of – 5 V. At the other extreme, an LM7924 produces an output of – 24 V.

69.3 Regulated Dual Supplies

By combining an LM78XX and an LM79XX, as shown in Fig. 69.2, we can regulate the output of a dual supply. The LM78XX regulates the positive output, and the LM79XX handles the negative output. The input capacitors prevent oscillations, and the output capacitors improve transient response. The two diodes ensure that both regulators can turn on under all operating conditions.

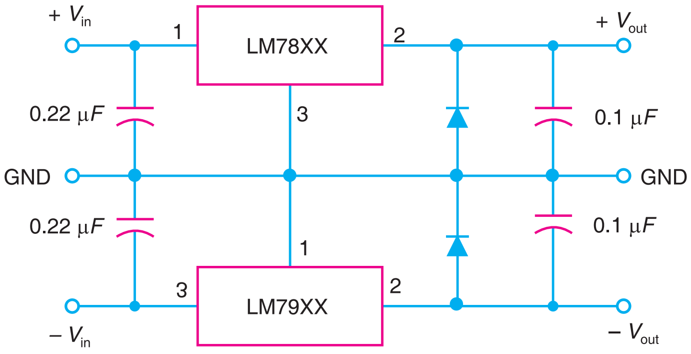

+ V_{in}
1
LM78XX
2
+ V_{out}
0.22 μF
3
0.1 μF
GND
GND
0.22 μF
0.1 μF
1
3
LM79XX
2
– V_{in}
– V_{out}

Fig. 69.2

69.4 Adjustable Voltage Regulator

LM 723 C is the general purpose adjustable voltage regulator. The output voltage is adjustable from 2 to 37 V.

We can understand the regulating function of the chip LM 723C by considering its internal circuit (Fig. 69.3).

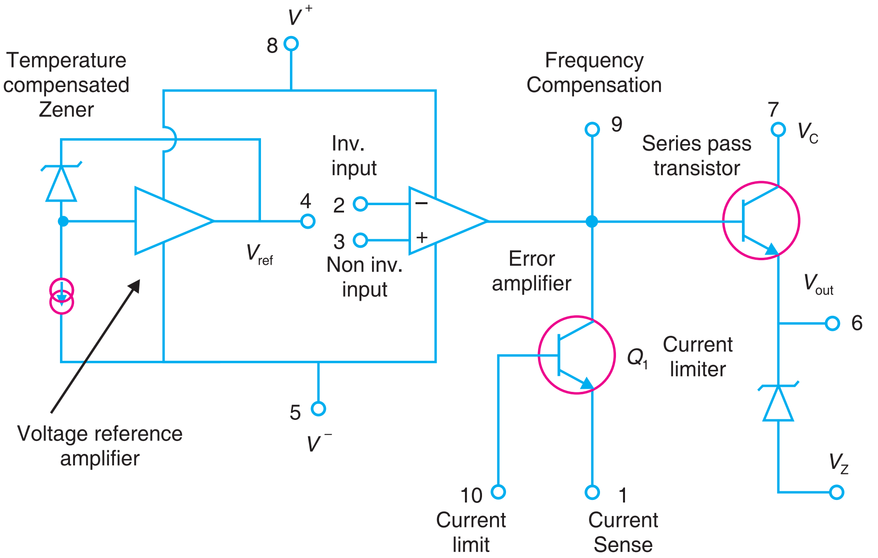
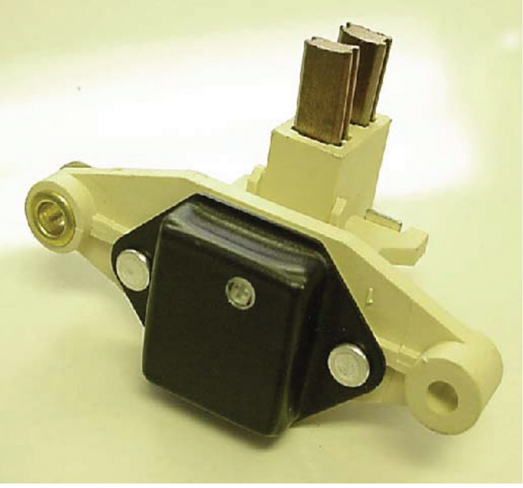

V^+
8
Temperature compensated Zener
Frequency Compensation
7
V_C
9
Series pass transistor
Inv. input
4
2
3
V_{ref}
Non inv. input
Error amplifier
V_{out}
6
Q_1
Current limiter
Voltage reference amplifier
5
V^-
V_Z
10
Current limit
1
Current Sense

Fig. 69.3

A reference voltage is developed across the zener diode which is temperature compensated. The reference amplifier acts as a buffer. So the constant reference voltage is available at its output. The control amplifier has two inputs, one inverting and the other non-inverting. This amplifier compares the reference with a fixed part of the feedback output voltage and the resultant error is amplified. This error voltage controls the series transistor such that the output voltage remains at a constant level. Transistor Q_1 is used for current limiting.



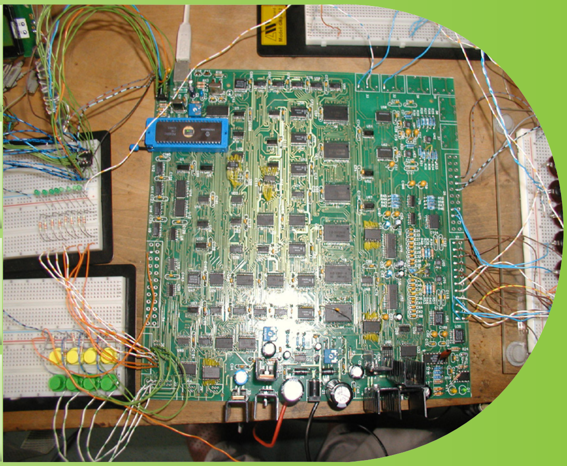
Adjustable Voltage Regulator.

The load current can be limited by providing a small resistance R_{SC} between the current limit and current sense terminals. The voltage across R_{SC} is used to bias the current limiting transistor, present inside the chip.

Hence, $V_{BE} = I_{\text{limit}} R_{SC}$, where V_{BE} is the bias voltage and I_{limit} is the limiting current.

$$R_{SC} = \frac{V_{BE}}{I_{\text{limit}}}.$$

CHAPTER
70



DIGITAL ELECTRONICS

AT A GLANCE

70.1	Decimal Number System	**70.2**	Binary Number System
70.3	Conversion of Binary Number into Decimal Number		
70.4	Conversion of Decimal Number into Binary Number		
70.5	Binary Addition	**70.6**	Binary Subtraction
70.7	Complement Method of Subtraction	**70.8**	Binary Multiplication
70.9	Binary Division	**70.10**	Octal Number System
70.11	Hexadecimal Number System	**70.12**	Boolean Algebra
70.13	Boolean Addition and Multiplication	**70.14**	Logic Expressions
70.15	Commutative Laws	**70.16**	Associative Laws
70.17	Distributive Law	**70.18**	Rules for Boolean Algebra
70.19	De Morgan's Theorems	**70.20**	Digital Logic Gates
70.21	The OR Gate	**70.22**	The AND Gate
70.23	The NOT Gate	**70.24**	Flip-Flop
70.25	Master-Slave Flip-Flop	**70.26**	Master-Slave JK Flip-Flop
70.27	Counters	**70.28**	The 7493 A Four-Bit Binary Counter
70.29	Shift Register	**70.30**	Serial in-Serial out Shift Register

NUMBER SYSTEMS

70.1 Decimal Number System

This is the frequently used number system in our daily life. It uses ten numerals 0, 1, 2, 3, 4, 5, 6, 7, 8 and 9. The *base* or the *radix* of the decimal system is 10. A number in decimal system is expressed in terms of the position or place values.

For example, the number 10523 is represented as

$10523 = 1 \times 10^4 + 0 \times 10^3 + 5 \times 10^2 + 2 \times 10^1 + 3 \times 10^0$.

The digit (3) has the position value 10^0 and is the *least significant digit* (*LSD*).

The digit (1) has the place value 10^4 and is the *most significant digit* (*MSD*).

Similarly, the number 2564.397 can be expressed as

$2564.397 = 2 \times 10^3 + 5 \times 10^2 + 6 \times 10^1 + 4 \times 10^0 + 3 \times 10^{-1} + 9 \times 10^{-2} + 7 \times 10^{-3}$.

That is the powers to the base 10 are numbered to the left of the decimal point starting with 0 and to the right of the decimal point starting with – 1.

70.2 Binary Number System

In a *binary system* of representations the *base* (or *radix*) is 2. It uses only two numerals 0 and 1. In a digital system there are only two possible states or conditions. For example, a situation may be True or False, a switch close or open, a voltage signal High or Low etc. These states or conditions are designated as 1 and 0 respectively. The binary digits 0 and 1 are termed as bits. Like the decimal system, the binary system also has a place or position value representation.

For example, the number 15 of decimal system is written in the binary system as 1111, since,

$$1111 = 1 \times 2^3 + 1 \times 2^2 + 1 \times 2^1 + 1 \times 2^0 = 15.$$

This can also be written as $1111_2 = 15_{10}$, the subscript indicating the number system. In the binary 1111, the bit 1 at the extreme left is the MSB (most significant bit) and the bit 1 at the extreme right is the LSB.

Similarly, the binary 1011.011 is written in the decimal system as 11.375, since

$$\begin{aligned} 1011.011 &= 1 \times 2^3 + 0 \times 2^2 + 1 \times 2^1 + 1 \times 2^0 + 0 \times 2^{-1} + 1 \times 2^{-2} + 1 \times 2^{-3} \\ &= 8+0+2+1+0+\frac{1}{4}+\frac{1}{8} = 11\cdot375 \end{aligned}$$

or $$1011.011_2 = 11.375_{10}$$

Thus, the place values of the bits in a binary number are given by ascending powers of 2 to the left of binary point starting from 0 and to the right of binary point in the descending powers of 2 starting from – 1.

70.3 Conversion of Binary Number into Decimal Number

(*i*) Conversion of integral binary numbers

EXAMPLE 1. *Convert* $(1010)_2$ *into its decimal equivalent.*

$$1010 = 1 \times 2^3 + 0 \times 2^2 + 1 \times 2^1 + 0 \times 2^0 = 8 + 0 + 2 + 0 = 10$$

$$\Rightarrow (1010)_2 = (10)_{10}$$

(*ii*) Conversion of fractional binary numbers

As an example, let us find the decimal equivalent of 0.1101.

$$\begin{aligned} 0.1101 &= 1 \times 2^{-1} + 1 \times 2^{-2} + 0 \times 2^{-3} + 1 \times 2^{-4} \\ &= \frac{1}{2}+\frac{1}{4}+0+\frac{1}{16} = 0.8125 \end{aligned}$$

$$\therefore \quad \underset{\substack{\downarrow \\ \text{Binary} \\ \text{Point}}}{0.1101_2} = \underset{\substack{\downarrow \\ \text{Decimal} \\ \text{Point}}}{0.8125_{10}}$$

(*iii*) Mixed Number

For mixed numbers, that is, having integer and fractional parts, each part is handled separately according to the following rules.

(*i*) The integer numbers are multiplied by their corresponding powers of base 2, *i.e.*, $2^0, 2^1, 2^2$, ... from left of the binary point and then added.

(*ii*) The fractional numbers are multiplied by their corresponding powers of base 2, *i.e.*, $2^{-1}, 2^{-2}, 2^{-3}$, ... from right of the binary point and then added.

Consider the binary number 1101.101. Its decimal equivalent is

$$1101.101 = 1 \times 2^3 + 1 \times 2^2 + 0 \times 2^1 + 1 \times 2^0 + 1 \times 2^{-1} + 0 \times 2^{-2} + 1 \times 2^{-3}$$

$$= 8+4+0+1+\frac{1}{2}+0+\frac{1}{8} = 13.625$$

Example 2. *Convert* $(1011.0101)_2$ *into its decimal equivalent.*

$$1011.0101 = 1 \times 2^0 + 1 \times 2^1 + 0 \times 2^2 + 1 \times 2^3 + 0 \times 2^{-1} + 1 \times 2^{-2} + 0 \times 2^{-3} + 1 \times 2^{-4}$$

$$= 1+2+0+8+0+\frac{1}{4}+0+\frac{1}{16}$$

$$= 11.3125$$

$$(1011.0101)_2 = (11.3125)_{10}$$

70.4 Conversion of Decimal Number into Binary Number

(*i*) Conversions of integral decimal numbers

The given decimal number is divided progressively by 2, until we get zero. The remainders, taken in the reverse order, give the binary number.

As an example let us convert the decimal 19 into its binary equivalent.

Successive divisions		**Remainders**	
2	19		
2	9	1	Top
2	4	1	
2	2	0	
2	1	0	
	0	1	Bottom

Reading the remainders from the bottom to the top, the binary equivalent of 19 is found to be 10011, or

$$19_{10} = 10011_2.$$

Example 3. *Convert* $(25)_{10}$ *into binary number.*

		Remainders
2	25	
2	12	1
2	6	0
2	3	0
2	1	1
	0	1

$$(25)_{10} = (11001)_2$$

(*ii*) Conversion of fractional-decimal numbers

The given decimal number is multiplied by 2 progressively. For each step that results in a 1 in the units place, transfer the 1 to the binary record and repeat the process with the fractional number. For each multiplication by 2 that results in a product less than unity, record a 0 in the binary number and carry on the process. The last step is reached if the fractional part is zero or it is terminated when the desired accuracy is attained. The carries are taken in the forward (top to bottom) direction to give the equivalent binary.

Let us convert 0.9125 into its binary equivalent.

$$\begin{aligned}
0.9125 \times 2 &= 1.8250 = 0.8250 \text{ with a carry } 1 \\
0.8250 \times 2 &= 1.6500 = 0.6500 \text{ with a carry } 1 \\
0.6500 \times 2 &= 1.3000 = 0.3000 \text{ with a carry } 1 \\
0.3000 \times 2 &= 0.6000 = 0.6000 \text{ with a carry } 0 \\
0.6000 \times 2 &= 1.2000 = 0.2000 \text{ with a carry } 1 \\
0.2000 \times 2 &= 0.4000 = 0.4000 \text{ with a carry } 0 \quad \downarrow
\end{aligned}$$

The process is terminated here to get an *approximate* result, namely, representation of 0.9125 by six binary digits.

$$\therefore \qquad 0.9125_{10} = 0.111010_2$$

The point in front of the binary is referred to as the *binary point*.

To find the binary equivalent of a decimal number like 35.625, we split the number into an integer of 35 and a fraction of 0.625. Then the binary equivalent of each part is obtained separately by methods discussed above.

First let us find out the binary equivalent of the integer part 35 by divide-by-two method.

$$\begin{aligned}
35 \div 2 &= 17 + 1 &&\text{remainder } 1 &&\text{Top} \\
17 \div 2 &= 8 + 1 &&\text{remainder } 1 &&\uparrow \\
8 \div 2 &= 4 + 0 &&\text{remainder } 0 \\
4 \div 2 &= 2 + 0 &&\text{remainder } 0 \\
2 \div 2 &= 1 + 0 &&\text{remainder } 0 \\
1 \div 2 &= 0 + 1 &&\text{remainder } 1 \\
\therefore \qquad 35_{10} &= 100011_2
\end{aligned}$$

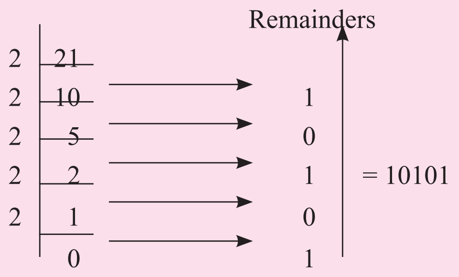
The binary equivalent of 0.625 is found by the multiply-by-two method,

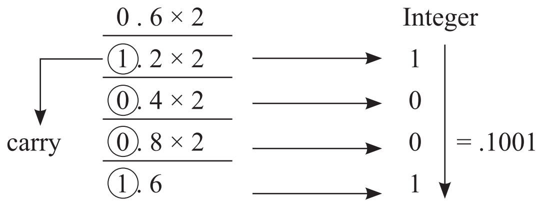
$$\begin{aligned}
0.625 \times 2 &= 1.25 &&= 1 + 0.25 &&\text{carry } 1 &&\text{Top} \\
0.25 \times 2 &= 0.50 &&= 0 + 0.50 &&\text{carry } 0 &&\downarrow \\
0.50 \times 2 &= 1.0 &&= 0 + 1 &&\text{carry } 1 &&\text{Bottom}
\end{aligned}$$

Taking the carrys from top to bottom

$$\begin{aligned}
0.625_{10} &= 101 \\
\therefore \qquad 35.625_{10} &= 100011.101_2
\end{aligned}$$

EXAMPLE 1. *Convert* $(21.6)_{10}$ *into binary number.*





∴ $(21.6)_{10} = (10101.1001)_2$

70.5 Binary Addition

The sum of two binary numbers is calculated by the same rules as in decimal, except that the digits of the sum in any significant position can be only 0 or 1. Any "carry" obtained in a given significant position is used by the pair of digits one significant position higher. Binary addition is with the four rules :

1. $0 + 0 = 0$
2. $0 + 1 = 1$
3. $1 + 0 = 1$
4. $1 + 1 = 10$ (one-zero, not ten)

The last rule is often written as $1 + 1 = 0$ with a carry of 1.

Example of addition of two binary numbers is shown below :

$$\begin{array}{rr} \text{augend :} & 101101 \\ \text{addend :} & +\,100111 \\ \hline \text{sum :} & 1010100 \end{array}$$

Consider **another example**

$$\begin{array}{cccccc} & 1 & 0 & 1 & 1 & (11) \\ + & 1 & 0 & 0 & 1 & (9) \\ \hline 1 & 0 & 1 & 0 & 0 & (20) \end{array}$$

Explanation. Addition is made column-wise.

70.6 Binary Subtraction

The *four rules* for binary subtraction are given below :

1. $0 - 0 = 0$
2. $1 - 0 = 1$
3. $1 - 1 = 0$
4. $10 - 1 = 1.$

The last rule indicates that when 1 is subtracted from 10_2 (= decimal 2), we get 1.

EXAMPLE 1.

$$\begin{array}{rr} \text{minuend :} & 101101 \\ \text{subtrahend :} & -\,100111 \\ \hline \text{difference :} & 000110 \end{array}$$

EXAMPLE 2. *Subtract* 1001 *from* 1110.

$$\begin{array}{rlrr} 1110 & \text{First column : (after borrow)} & 10-1\ =\ 1 & 14 \\ (-)\ 1001 & \text{2nd column : ....} & 0-0\ =\ 0 & (-)\ 9 \\ \hline 0101 & \text{3rd column : ....} & 1-0\ =\ 1 & \hline 5 \\ \hline & \text{4th column : ....} & 1-1\ =\ 0 & \end{array}$$

70.7 Complement Method of Subtraction

(*a*) 1's Complement. The 1's complement of a binary number is the number which is obtained by changing its each 0 into a 1 and each 1 into a 0. For example, 1's complement of the binary number 1011 is 0100.

Subtraction by 1's complement :

(*i*) The 1's complement of the subtrahend is computed.

(*ii*) The 1's complement is added to the minuend.

(*iii*) If there is no end-around carry, then the answer is recomplemented and a negative sign is attached to it.

(*iv*) If there is an end-around carry of 1, remove the carry 1 from the addition and add it to the remainder. In this case no recomplementing is necessary.

EXAMPLE 1. *Let us subtract* 1101_2 *from* 1010_2.

1's Complement Method **Conventional Method**

$$\begin{array}{r} 1010 \\ +\,0010 \\ \hline \boxed{}1100 \\ \uparrow \\ \text{No carry} \end{array} \qquad (1\text{'s complement of } 1101) \qquad \begin{array}{r} 1010 \\ -\,1101 \\ \hline -\,0011 \end{array}$$

Since there is no end-around carry (EAC) in this case, we write 1's complement of 1100 and attach a minus sign to it. Thus

Final answer → – 0011.

EXAMPLE 2. *Consider the subtraction*

$$\begin{array}{rl} 1010 & \quad \text{Minuend} \\ -\,1001 & \quad \text{Subtrahend} \end{array}$$

1's complement of the subtrahend is 0110

$$\therefore \quad \begin{array}{r} 1010 \\ -1001 \\ \hline \end{array} \;\Rightarrow\; \begin{array}{r} 1010 \\ +\,0110 \\ \hline 10000 \\ \uparrow \\ \text{carry} \end{array} \qquad \Rightarrow \begin{array}{r} 0000 \\ +\quad 1 \\ \hline 0001 \end{array}$$

$$\therefore \quad \begin{array}{r} 1010 \\ -\,1001 \\ \hline 0001 \end{array}$$

***(b)* 2's complement.** The 2's complement of a binary number is obtained by adding 1 to its 1's complement.

Subtraction by 2's complement

(*i*) Find the 2's complement of the subtrahend.
(*ii*) Add this complement to the minuend.
(*iii*) Drop the carry in the last position.
(*iv*) If the carry in the last position is 1, the result is positive.
(*v*) If there is no 1 carry in the last position, determine the 2's complement of the result of subtraction and attach a minus sign to it. Thus, the answer of subtraction is negative.

The following examples illustrate the procedure.

(*i*) Let us subtract 10001_2 from 10011_2.

$$\begin{array}{rcr} \textbf{2's Complement Method} & & \textbf{Conventional Method} \\ 10011 & & 10011 \\ +\,01111 & \text{(2's complement of 10001)} & -\,10001 \\ \cline{1-1}\cline{3-3} [1]\,00010 & & 00010 \\ \downarrow & & \end{array}$$

Drop the carry.

Therefore, *final answer is* : 00010.

(*ii*) Let us subtract 1101_2 from 1010_2.

$$\begin{array}{rl} 1010 & \quad\text{Minuend} \\ -\,1101 & \quad\text{Subtrahend} \end{array}$$

2's complement of subtraction = 1's complement + 1 = 0010 + 1 = 0011

$$\begin{array}{rl} 1010 & \\ +\,0011 & \quad\text{(2's complement of 1101)} \\ \hline \boxed{}\;1101 & \\ \uparrow & \end{array}$$

No carry

Thus, there is no 1 carry in the last position. Hence, we determine 2's complement of 1101 which is

$$0010 + 1 = 0011.$$

After attaching the minus sign, the final answer is – 0011.

70.8 Binary Multiplication

The following are the rules for binary multiplication.

(*i*) $0 \times 0 = 0$
(*ii*) $0 \times 1 = 0$
(*iii*) $1 \times 0 = 0$
(*iv*) $1 \times 1 = 1$

Binary multiplication is carried out as in decimal system.

EXAMPLE 1. *Multiply* 10110 *by* 110.

$$\begin{array}{r} \textbf{10110} \times \textbf{110} \\ 00000 \\ 10110 \\ 10110 \\ \hline 10000100 \end{array} \qquad\qquad \begin{array}{c} \textbf{Verification} \\ \hline 22 \times 6 \\ 132 \end{array}$$

∴ The result is 10000100.

70.9 Binary Division

Binary division is done as in decimal system.

EXAMPLE 2. *Divide* 1111 by 110.

$$\begin{array}{r|l} & 10 \\ \cline{2-2} 110 & 1111 \\ \cline{2-2} & 110 \\ & 11 \end{array} \qquad\qquad \begin{array}{l} \text{Verification} \\ 15 \div 6 \text{ gives} \end{array}$$

$$\begin{array}{r|l} & 2 \\ \cline{2-2} 6 & 15 \\ \cline{2-2} & 12 \\ \cline{2-2} & 3 \end{array}$$

∴ Quotient : 10

Remainder : 11

Quotient : 2

Remainder : 3

70.10 Octal Number System

The radix or base of this system is 8. It uses the eight numerals 0, 1, 2, 3, 4, 5, 6, 7. For counting beyond 7, 2-digit combinations are formed taking the second digit followed by the first, then the second digit followed by the second and so on. For example, after 7, the next number in the octal system is 10 (second digit followed by the first), then 11 (second digit followed by the second). The position value of each digit is in ascending powers of 8 for integers and descending powers of 8 for fractions as shown below :

$$\leftarrow 8^3\ \ 8^2\ \ 8^1\ \ 8^0\quad \underset{\substack{\uparrow \\ \text{octal point}}}{\bullet} \quad 8^{-1}\ \ 8^{-2}\ \ 8^{-3} \rightarrow$$

For example, 425.350 in the octal system can be expressed in decimal system in the following way.

$$\begin{aligned} 425.350 &= 4 \times 8^2 + 2 \times 8^1 + 5 \times 8^0 + 3 \times 8^{-1} + 5 \times 8^{-2} + 0 \times 8^{-3} \\ &= 4 \times 64 + 2 \times 8 + 5 \times 1 + 3 \times \frac{1}{8} + 5 \times \frac{1}{64} \\ &= 256 + 16 + 5 + 0.375 + 0.078 = 277.453_{10} \end{aligned}$$

70.11 Hexadecimal Number System

In the hexadecimal system the radix or base is 16. It uses the sixteen numerals

0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, F.

Thus, A represents 10, B represents 11, ... F represents 15. After reaching F, two digit combinations are formed taking the second digit followed by the first, then the second followed by the second and so on. For example, in the hexadecimal system 10 (second digit followed by the first) represents 16 in the decimal system. The position value for each digit is in ascending powers of 16 for integers and descending powers of 16 for fractions.

Examples of Hexadecimal to Decimal Conversion

(*i*) $3\,C\,8_{16} = 3 \times 16^2 + 12 \times 16^1 + 8 \times 16^0$

$= 768 + 192 + 8 = 968_{10}$

(*ii*) $E\,5\,F\,8_{16} = 14 \times 16^3 + 5 \times 16^2 + 15 \times 16^1 + 8 \times 16^0$

$= 57344 + 1280 + 240 + 8 = 58{,}872_{10}$

BOOLEAN ALGEBRA

70.12 Boolean Algebra

The algebra of logic prominently used in the operation of computer devices is the algebra developed by George Boole. It is a binary or two-valued logic, *i.e.*, it permits only two values or states for its variables. These two states are 'true' and 'false' in logic but are represented by 'on' and 'off' states of electronic circuits.

The two variables of the Boolean algebra are usually represented by 0 and 1. Hence, every variable is either a 0 or a 1. There are no negative or fractional numbers. Logically, we may write :

If $X = 0$ then $X \neq 1$

And if $X = 1$ then $X \neq 0$.

Boolean algebra uses only three operations on its variables. These operations are:

(*i*) The OR addition represented by a + (plus) sign.

(*ii*) The AND multiplication represented by a × (cross), or a . (dot) sign.

We will normally write the adjacent letters without the dot, such as *AB*.

(*iii*) The NOT operation represented by a bar over a variable.

70.13 Boolean Addition and Multiplication

Addition in Boolean algebra involves variables having values of either a binary 1 or a binary 0. The basic rules for Boolean addition are as follows :

$$0 + 0 = 0$$
$$0 + 1 = 1$$
$$1 + 0 = 1$$
$$1 + 1 = 1$$

In the application of Boolean algebra to logic circuits, *Boolean addition is the same as the OR*.

The basic rules for Boolean multiplication are :

$$0 \, . \, 0 = 0$$
$$0 \, . \, 1 = 0$$
$$1 \, . \, 0 = 0$$
$$1 \, . \, 1 = 1$$

Boolean multiplication is the same as the AND.

70.14 Logic Expressions

NOT

The operation of an invertor (NOT circuit) can be expressed with symbols as follows : If the input variable is called A and the output variable is called X, then $X = \overline{A}$. This expression states that the output is the complement of the input, so that if $A = 0$, then $X = 1$, and if $A = 1$, then $X = 0$ (Fig. 70.1).

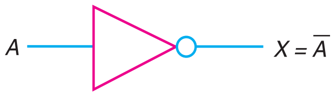


Fig. 70.1

(*i*) $\overline{0} = 1$ (*ii*) $\overline{1} = 0$

AND

The operation of a two-input AND gate can be expressed in equation form as follows : If one input variable is A, the other input variable is B, and the output variable is X, then the Boolean expression for this basic gate function is $X = AB$ (Fig. 70.2).

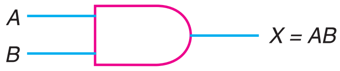


Fig. 70.2

OR

The operation of a two-input OR gate can be expressed in equation form as follows : If one input is A, the other input is B, and the output is X, then the Boolean expression is $X = A + B$ (Fig. 70.3).

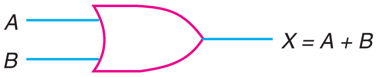


Fig. 70.3

70.15 Commutative Laws

(*i*) The commutative law of addition for two variables is written algebraically as

$$A + B = B + A$$

This states that the order in which the variables are ORed makes no difference.

(*ii*) The commutative law of multiplication of two variables is

$$AB = BA$$

This states that the order in which the variables are ANDed makes no difference.

70.16 Associative Laws

(*i*) The associative law of addition is stated as follows for three variables :

$$A + (B + C) = (A + B) + C$$

This law states that in the ORing of several variables, the result is the same regardless of the grouping of the variables.

(*ii*) The associative law of multiplication is stated as follows for three variables :

$$A\,(BC) = (AB)\,C$$

This law tells us that it makes no difference in what order the variables are grouped when ANDing several variables.

70.17 Distributive Law

The distributive law is written for three variables as follows :

$$A\;(B + C) = AB + AC$$

This law states that ORing several variables and ANDing the result with a single variable is equivalent to ANDing the single variable with each of the several variables and then ORing the products.

70.18 Rules for Boolean Algebra

Table 70.1 lists *Basic rules of Boolean algebra.*

Table 70.1

1. $A + 0 = A$
2. $A + 1 = 1$
3. $A \,.\, 0 = 0$
4. $A \,.\, 1 = A$
5. $A + A = A$
6. $A + \overline{A} = 1$
7. $A \,.\, A = \mathrm{A}$
8. $A \,.\, \overline{A} = 0$
9. $\overline{\overline{A}} = A$
10. $A + AB = A$
11. $A + \overline{A}B = A + B$
12. $(A + B)\,(A + C) = \mathrm{A} + BC$

Rule 1. A variable ORed with a 0 is equal to the value of the variable ($A + 0 = A$) (Fig. 70.4).

Similarly, the other rules are proved.

70.19 De Morgan's Theorems

Theorem 1. *The complement of the product of two variables is equal to the sum of complements of the variables, i.e.,*

$$\overline{AB} = \overline{A} + \overline{B}$$

The complement of two variables ANDed is the same as the OR of the complements of each individual variable (Fig. 70.5).

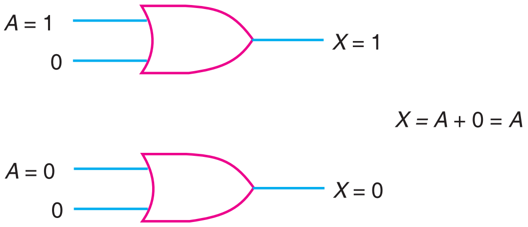


Fig. 70.4

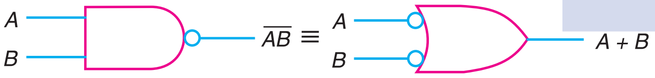


Fig. 70.5

Theorem 2. *The complement of the sum of two variables is equal to the product of complements of the variables, i.e.,*

$$\overline{A + B} = \overline{A}\,\overline{B}$$

The complement of two variables ORed is the same as the AND of the complements of each individual variable (Fig. 70.6).

Truth Table

A	B	$\overline{AB}$	$\overline{A} + \overline{B}$
0	0	1	1
0	1	1	1
1	0	1	1
1	1	0	0

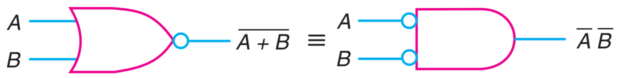


Fig. 70.6

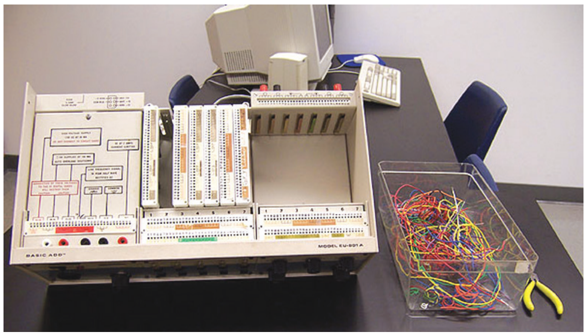
Truth Table

A	B	$\overline{A+B}$	$\overline{A}\,\overline{B}$
0	0	1	1
0	1	0	0
1	0	0	0
1	1	0	0

LOGIC GATES

70.20 Digital Logic Gates

A logic gate is an electronic circuit which performs logic functions or takes a logic decision. It has one output and one or more inputs. Logic gates are the building blocks of the digital systems. They work on the logical algebra developed by George Boole. The Boolean operations namely 'OR' operation, 'AND' operation and 'NOT' operation are implemented by three logic gates called 'OR' gate, 'AND' gate and 'NOT' gate.

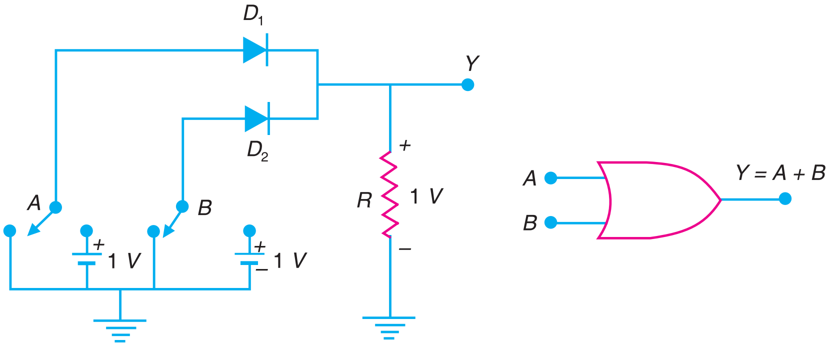


Binary Number and Logic Gates.

70.21 The OR Gate

It implements Boolean 'OR' operation. An 'OR' gate has two (or) more inputs and one output. *An* 'OR' *gate is a logic gate whose output is* '1' *state if any or all the inputs are in* '1' *state.*

Fig. 70.7 (*a*) shows a two input OR gate using two ideal diodes D_1 and D_2. Here A and B represent the two inputs and Y the output. R represents the output load resistor. Fig. 70.7 (*b*) gives the symbolic representation of the OR gate.



(*a*) (*b*)

Fig. 70.7

Suppose that the two input voltages are either 0 or 1 V.

Case (*i*). When $A = 0$; $B = 0$, both the input voltages are zero. So both the diodes D_1 and D_2 do not conduct. Hence, current through R is zero and so output voltage is zero, *i.e.*, $Y = 0$.

Case (*ii*). When $A = 1$; $B = 0$, the diode D_1 is forward biased and so it conducts. D_2 does not conduct. Since D_1 is conducting, current flows through R. Hence, there is an output voltage, *i.e.*, $Y = 1$.

Case (*iii*). When $A = 0$; $B = 1$, the diode D_2 is forward biased and so it conducts. D_1 does not conduct. Since D_2 is conducting, current flows through R. Hence, there is an output voltage, *i.e.*, $Y = 1$.

Case (*iv*). When $A = 1$; $B = 1$, both the diodes D_1 and D_2 are forward biased. Hence, both are conducting. So current flows through R and there is an output voltage, *i.e.*, $Y = 1$.

The logic operation of the OR gate can be summarised in a tabular form known as Truth Table. A truth table may be defined as a table which gives the output state for all possible input combinations.

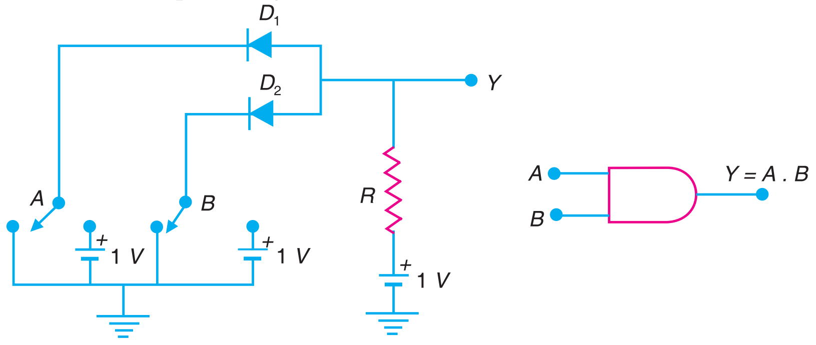
Truth Table

A	B	$Y = A + B$
0	0	0
1	0	1
0	1	1
1	1	1

70.22 The AND Gate

An "AND" gate implements Boolean 'AND' operation. It has two or more inputs and one output. *An 'AND' gate is a logic gate whose output is* '1', *if and only if all the inputs are in* '1' *state.*

Fig. 70.8 (*a*) *shows a two-input AND gate using two ideal diodes and Fig.* 70.8 (*b*) *gives its symbolic representation.* Here A and B represent the two inputs and Y the output. R represents the output load resistor. The two input voltages are assumed to be either 0 or 1 V.



(*a*) (b)

Fig. 70.8

(*i*) When both $A = 0$ V and $B = 0$ V, the inputs are short circuited to ground. The 1 V battery in the output side biases the diodes D_1 and D_2 in the forward direction. Hence, both diodes conduct. The output is also shorted to ground through the diodes. Thus, the output $Y = 0$ V.

(*ii*) When $A = 1$ V and $B = 0$ V, diode D_2 conducts and the output is short circuited to ground through this diode. Therefore, output $Y = 0$ V.

(*iii*) When $A = 0$ V and $B = 1$ V, diode D_1 conducts and the output is short circuited to ground through this diode. Therefore, the output $Y = 0$ V.

(*iv*) When both $A = 1$ V and $B = 1$ V, neither D_1 nor D_2 conducts. No current, therefore, flows through R and the output $Y = 1$ V.

The above results are tabulated in the Truth table.

Truth Table

A	B	$Y = A \cdot B$
0	0	0
1	0	0
0	1	0
1	1	1

70.23 The NOT Gate

The NOT circuit has only one input and one output. It inverts the polarity of a pulse applied to it. If the input is 1, the output is 0 and when the input is 0 the output is 1. That is, the NOT gate inverts the input. A 'NOT' gate using transistor is shown in Fig. 70.9 (*a*) and its symbol in Fig. 70.9 (*b*).

(*i*) When no signal is applied at the input, *i.e.*, $A = 0$ V, the transistor is cut OFF, making the collector current zero. Thus, the potential drop across R is zero. The supply voltage of V_{cc} appears at the output terminal. Hence, the output $Y = V_{cc}$.

Thus, *when input is low, output is high*.

(*ii*) When a positive pulse is applied to A, *i.e.*, $A = 1$, the transistor conducts (fully ON) drawing maximum collector current. Hence, whole of V_{cc} drops across R and output $Y = 0$ V. *Thus, when input is high, the output is held at a low value.*

The results are tabulated in the Truth table.

Truth Table for NOT Gate

Input	*Output*
A	$Y = \overline{A}$
0 1	1 0

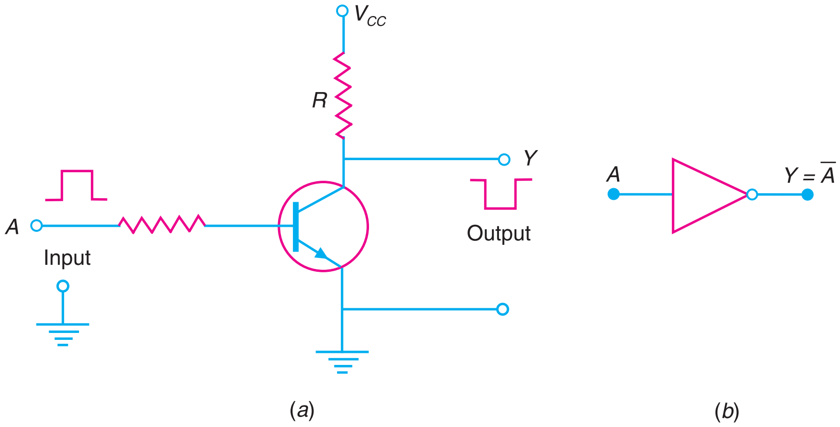
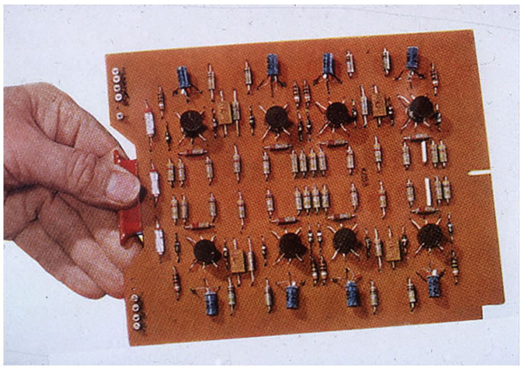


Fig. 70.9

FLIP FLOPS

70.24 Flip-Flop



Flip-Flop Circuits.

A 'Flip-Flop' is a bistable multivibrator which has two stable states, namely '1' and '0'. The circuit remains in the same state either in '1' state (or) in '0' state, unless it is changed by applying a triggering pulse. This property of the circuit is called as 'memory', *i.e.*, it can store one bit digital information. Since this information is locked in this circuit, it is also called as "latch".

A flip-flop circuit is a special electronic circuit with two output signals Q and $\overline{Q}$; the Q output being known as the true output and $\overline{Q}$ being complementary or false output (Fig. 70.10). These output signals assume one of two voltage levels that are complementary. If Q is high then $\overline{Q}$ is low and *vice-versa*. The output remains stable till it is changed by the input signal. The two output stable conditions are $Q = 1, \overline{Q} = 0$, and $\overline{Q} = 0$, $Q = 1$, and these are called states. $Q = 1$ is called the Set state or 1 state and $Q = 0$ is called the Reset state or 0 state.

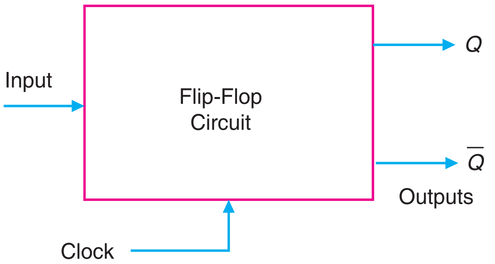


Fig. 70.10

70.25 Master-Slave Flip-Flop

A *master-slave flip-flop* is constructed from two separate flip-flops. One circuit serves as a master and the other as a slave. Fig 70.11 shows the logic diagram of an *RS* master-slave flip-flop. It consists of a master flip-flop, a slave flip-flop, and an inverter.

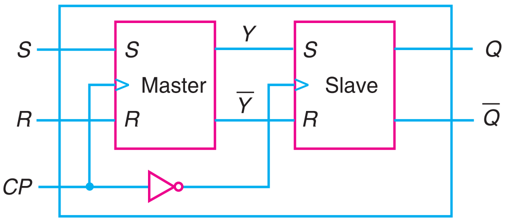


Fig. 70.11

When clock pulse *CP* is 0, the output of the inverter is 1. Since the clock input of the slave is 1, the flip-flop is enabled and output Q is equal to Y, while $\overline{Q}$ is equal to $\overline{Y}$. The master flip-flop is disabled because $CP = 0$. When the pulse becomes 1, the information then at the external R and S inputs is transmitted to the master flip-flop. The slave flip-flop, however, is isolated as long as the pulse is at its 1 level, because the output of the inverter is 0. When the pulse returns to 0, the master flip-flop is isolated, which prevents the external inputs from affecting it. The slave flip-flop then goes to the same state as the master flip-flop.

70.26 Master-Slave JK Flip-Flop

The circuit diagram is shown in Fig. 70.12. The clocks to the Master and Slave flip-flops are complementary to each other. It consists of two flip-flops. Gates 1 through 4 form the master flip-flop. Gates 5 through 8 form the slave flip-flop. The information present at the J and K inputs is transmitted to the master flip-flop on the positive edge of a clock pulse. It is held there until the negative edge of the clock pulse occurs, after which it is allowed to pass through to the slave flip-flop. The clock input is normally 0, which keeps the outputs of gates 1 and 2 at the 1 level. This prevents the J and K inputs from affecting the master flip-flop.

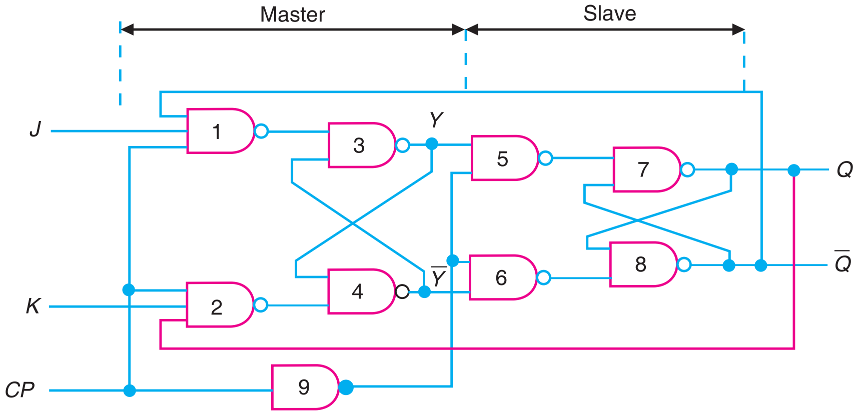


Fig. 70.12

The slave flip-flop is a clocked *RS* type, with the master flip-flop supplying the inputs and the clock input being inverted by gate 9. When the clock is 0, the output of gate 9 is 1. So, output $Q = Y$, and $\overline{Q} = \overline{Y}$.

When the positive edge of a clock pulse occurs, the master flip-flop is affected and may switch states. The slave flip-flop is isolated as long as the clock is at the 1 level, because the output of gate 9 provides a 1 to both inputs of the NAND basic flip-flop of gates 7 and 8. When the clock input returns to 0, the master flip-flop is isolated from the *J* and *K* inputs and the slave flip-flop goes to the same state as the master flip-flop.

70.27 Counters

In digital systems possibly the most widely used block is a counter. Counter is an instrument used for measuring time. A counter, which is driven by a clock can be used to count the number of clock pulses. There are two types of counters :

(*i*) Asynchronous (or) Binary ripple counters.

(*ii*) Synchronous counter (parallel counter).

In asynchronous counters, the clock pulses (CK) are applied to the first Flip-Flop and the output of the first Flip-Flop is used as CK for the next Flip-Flop and so on. In asynchronous counters, the FFs are connected in series and each FF takes its own propagation time. So the net output occurs only after the total propagation time. Hence, its speed is limited.

In synchronous counters (or parallel counters), the CKs are applied simultaneously to all the FFs. So the net propagation time becomes less and hence its speed is increased.

70.28 The 7493 A Four-Bit Binary Counter

It is an integrated circuit asynchronous counter. The logic diagram is shown in Fig. 70.13. Pin numbers are in parentheses, and all J-K inputs are internally connected HIGH.

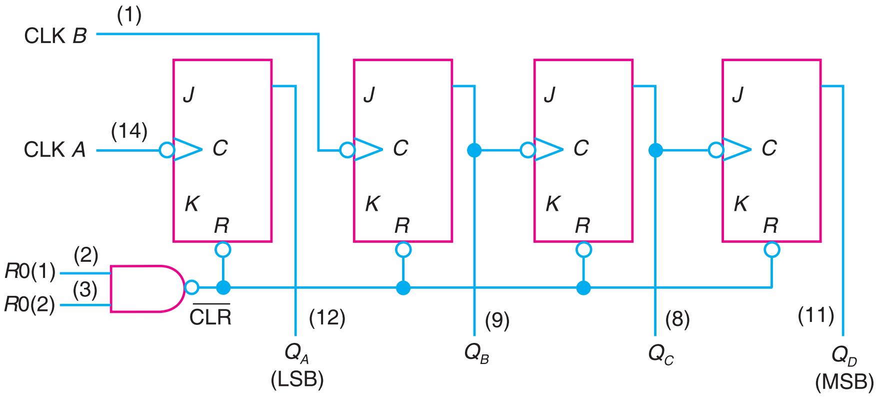


Fig. 70.13

This device actually consists of a single flip-flop and a three-bit asynchronous counter. This arrangement is for flexibility. It can be used as a divide-by-2 device using only the single flip-flop, or it can be used as a modulus-8 counter using only the three-bit counter portion. This device also provides gated reset inputs, *R*0 (1) and *R*0 (2). When both of these inputs are HIGH, the counter is RESET to the 0000 state by CLR.

Additionally, the 7493 A can be used as a four-bit modulus-16 counter (counts 0 through 15) by connecting the Q_A output to the CLKB input as shown in Fig. 70.14 (*a*).

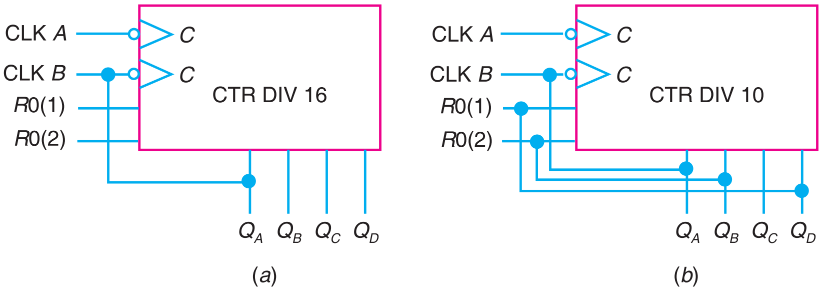

CLK A
CLK B
R0(1)
R0(2)
C
C
CTR DIV 16
Q_A Q_B Q_C Q_D
(a)
CLK A
CLK B
R0(1)
R0(2)
C
C
CTR DIV 10
Q_A Q_B Q_C Q_D
(b)

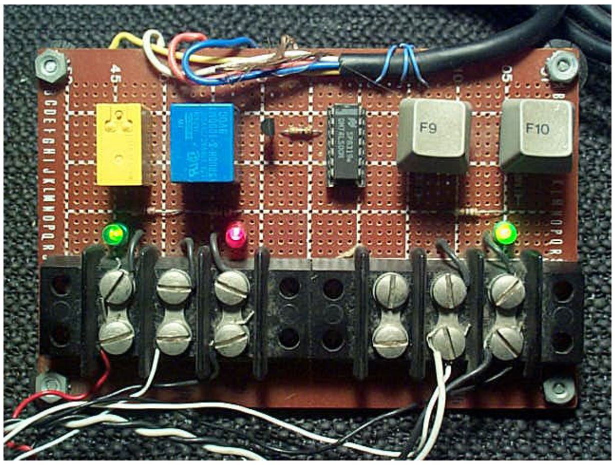
Fig. 70.14

It can also be configured as a decade counter (counts 0 through 9) with asynchronous recycling by using the gated reset inputs for partial decoding of count 10_{10}, as shown in Fig. 70.14 (*b*).

70.29 Shift Register

A register is a logical block in a digital system which can store the binary signal for one clock period. Registers are used extensively in arithmetic operations, counters and other digital systems to store binary data temporarily. We have seen above that a flip-flop can store one bit of data by assuming either a 1 or a 0 state. The simple single FF, therefore, is a one bit register. If a larger number of bits, say *N*, have to be stored then *N* flip-flops are required. The array of flip-flops with a number of bits stored in them will be a *N* bit register.

Sometimes it is required to shift the bits from one flip-flop to another. A shift register consists of group of flip-flops which are used to shift-in or shift-out (or both) the bits in a register. A shift register may have provisions to (*i*) enter data serially, parallely or both and (*ii*) shift the data to the right or to the left.



Shift Register.

70.30 Serial in-Serial out Shift Register

The serial in-serial out shift register constructed, using JK flip-flops, having clock inputs that are sensitive to negative clock transitions is shown in Fig. 70.15. The operation of the register is explained below.

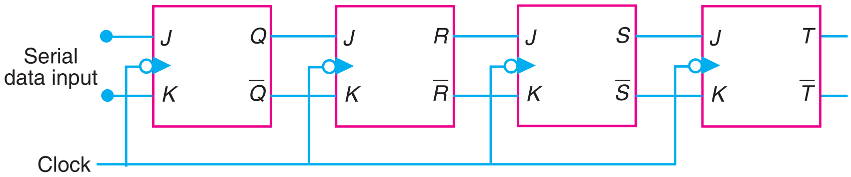

Serial
data input
Clock
J
K
Q
$\overline{Q}$
J
K
R
$\overline{R}$
J
K
S
$\overline{S}$
J
K
T
$\overline{T}$

Fig. 70.15

Consider four time instances t_1, t_2, t_3 and t_4 (Fig. 70.16).

(*i*) At the time instance t_1, all the flip-flops are reset making all J inputs low and K inputs high. The 0 in S is shifted to T. The 0 in R to S, the 0 in Q to R and the 0 at the input is shifted to Q. Thus, the output of the flip-flop at t_1 will be '0000'.

(*ii*) At the instant t_2, all the flip-flops have 0s. Thus, the 0 in S is shifted to T. 0 in R is shifted to S, 0 in Q is shifted to R and the 0 in data input is shifted to Q. Thus, the output of the flip-flop at the instant t_2 will be 0000.

(*iii*) At the instant t_3, all the flip-flops have 0s. The 0 in S is shifted to T, 0 in R is shifted to S and 0 in Q is shifted to R. However, the '1' in the data input is now shifted to Q. Thus the flip-flop output at the instant t_3 will be 1000.

(*iv*) At the instant t_4, the 0 in S is shifted to T, the 0 in R is shifted to S, the 1 in Q is shifted to R and the 0 in the data input shifts to Q. Thus, the flip-flop output at the instant t_4 will be 0100. Thus, a four bit data is shifted in a serial fashion.

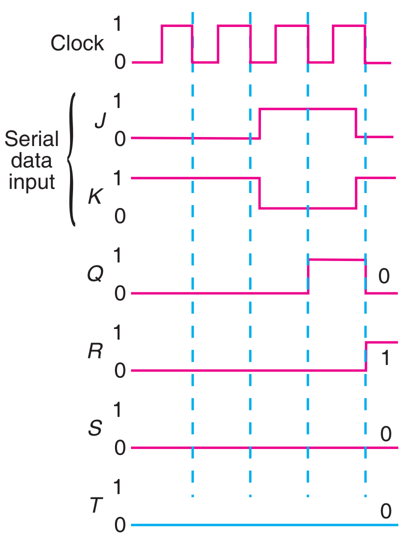


Fig. 70.16

EXERCISE

1. What is binary number system? How does it differ from decimal number system? Why is the binary number system used in computers?
2. What are binary numbers? Explain with illustrations the methods of conversion from decimal to binary and binary to decimal numbers.
3. Write four basic rules for adding binary digits. Give the truth table for binary addition.
4. State the rules for binary subtraction. Explain 1's complement and 2's complement method with examples.
5. What are octal and hexadecimal number systems?
6. What is Boolean algebra? Mention its unique feature. Discuss the fundamental laws of Boolean algebra. [*Karnatak University, Dharwad, Oct. 93*]
7. Write and explain the Boolean expressions for OR, AND and NOT circuits.
8. State and prove De Morgan's theorems.
9. (*a*) Draw logic symbols and truth tables of AND, OR, NOT logic gates. Explain its operation for two-input condition.

   (*b*) Draw the circuits of two-input AND, OR gates. Explain these two circuits. [*April 95, KUD*]
10. Explain the operation of a NOT circuit. Give its truth table and logic symbol.
11. Discuss the operation of a JK master-slave flip-flop with necessary circuit.
12. Explain the operation of a "Four Bit Binary Counter".
13. Define a register. What is a shift register? Explain the operation of a Serial in-Serial out shift register.
14. Solve the Boolean expression $X = A + B$ for the following inputs:

    (*i*) $A = 0, B = 0$; (*ii*) $A = 1, B = 0$; (*iii*) $A = 1, B = 1$.

Sol. (*i*) When $A = 0, B = 0, X = 0 + 0 = 0$.

(*ii*) When $A = 1, B = 0, X = 1 + 0 = 1$.

(*iii*) When $A = 1, B = 1, X = 1 + 1 = 1$.

15. Prove the following Boolean identities.

    (*i*) $AC + ABC = AC$

    (*ii*) $A + \overline{A}B = A + B$

Sol. (*i*) $AC + ABC = AC(1 + B) = AC$ ($\because$ $1 + B = 1$)

(*ii*) $A + \overline{A}B = A.1 + \overline{A}B$ [Law (4)]

$= A(1 + B) + \overline{A}B$ [Law (2)]

$= A.1 + AB + \overline{A}B$ (Distributive law)

$= A + B\,(A + \overline{A})$ (commutative law and distributive law)

$= A + B.1$ [Law (6)]

$= A + B.$

$\therefore$ $A + \overline{A}B = A + \text{B}.$

16. Reduce $AB + ABC + \overline{A}B + A\overline{B}C$ using laws of Boolean algebra.

Sol. $AB + ABC + \overline{A}B + A\overline{B}C$

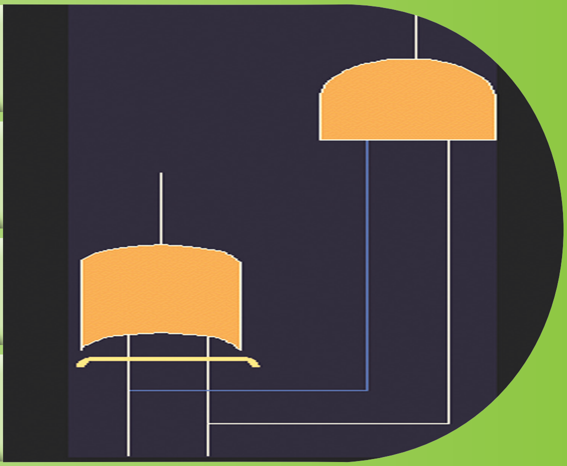
$= AB + \overline{A}B + ABC + A\overline{B}C$ [commutative law]

$= B(A + \overline{A}) + AC(B + \overline{B})$ [Distributive law and commutative Law]

$= B.1 + AC.1$ [Law (6)]

$= B + AC$ [Law (4)]

$\therefore$ $AB + ABC + \overline{A}B + A\overline{B}C = B + AC$.



CHAPTER

71

HALF-ADDER AND FULL ADDER

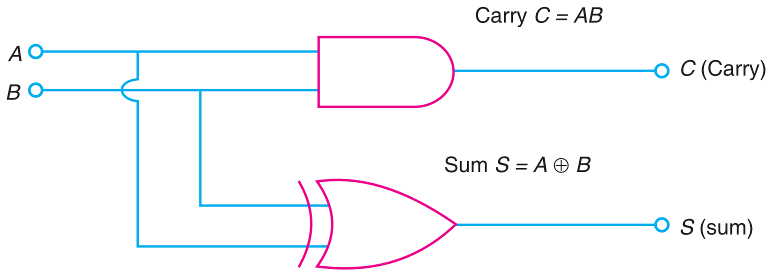
AT A GLANCE

71.1	Half-Adder	**71.2**	Full Adder
71.3	Parallel Binary Adder		

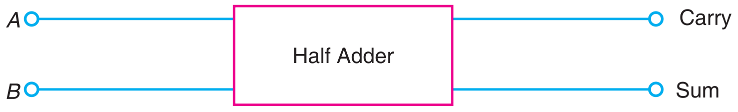
71.1 Half-Adder

A logic circuit that adds two bits producing a sum and a carry to be used in the next higher position is called a *half-adder*.

Fig. 71.1 (*a*) shows the circuit of a half-adder. It consists of an EXCLUSIVE OR gate and an AND gate. The output of the exclusive OR gate is called the *SUM*, while the output of the AND gate is called the CARRY. Fig. 71.1 (*b*) shows the symbol of a half-adder.


Carry $C = AB$
A
B
C (Carry)
Sum $S = A \oplus B$
S (sum)

(*a*) Half Adder


A
B
Half Adder
Carry
Sum

(*b*)

Fig. 71.1

The two inputs A and B represent the bits to be added.

S and C are the two outputs.

S represents the output of EXCLUSIVE OR gate.

C represents the CARRY bit. It is the output of AND gate.

With two inputs A and B, there are four distinct cases.

(I) When $A = 0$ and $B = 0$,

Carry $C = AB = 0$

Sum $S = A \oplus B = 0 \oplus 0 = 0$.

(II) When $A = 0$ and $B = 1$,

Carry $C = AB = 0.1 = 0$

Sum $S = A \oplus B = 0 \oplus 1 = 1$.

(III) When $A = 1$ and $B = 0$,

Carry $C = AB = 1.0 = 0$

Sum $S = A \oplus B = 1 \oplus 0 = 1$.

(IV) When $A = 1$ and $B = 1$,

Carry $C = AB = 1.1 = 1$

Sum $S = A \oplus B = 1 \oplus 1 = 0$.

The truth table for the Half-Adder operation is given in Table 1.

TRUTH TABLE 1

Input		*Output*	
A	*B*	*Sum* (*S*)	*Carry* (*C*)
0	0	0	0
0	1	1	0
1	0	1	0
1	1	0	1

The sum column represents XOR operation.

The carry column represents the AND operation.

The Boolean expressions for half-adder operation may be written as,

$$\text{Sum } S = A \oplus B$$

$$\text{Carry } C = AB.$$

71.2 Full Adder

We define full-adder as a logic circuit that adds three bits-two bits to be added and a carry bit from previous addition, which results in a sum and a carry.

The half-adder accomplishes the first step in the addition of bits. Two half-adders are combined into a *full adder*, in order to add the carry bits to the bit sum. Let us illustrate this operation by an example.

$$\begin{array}{cccc} & 1 & 1 & 1 \\ + & 1 & 0 & 1 \\ \hline & 11 & 0 & 0 \end{array}$$

In the least significant column,

$$1 + 1 = 10, \text{Sum} = 0, \text{Carry} = 1.$$

Thus for this operation, we need a half-adder.

In the next column, we must add 3 bits because of carry,

$$1 + 0 + 1 = 10, \text{Sum} = 0, \text{Carry} = 1.$$

By connecting two half-adders and an OR gate, we get a full adder (Fig. 71.2).

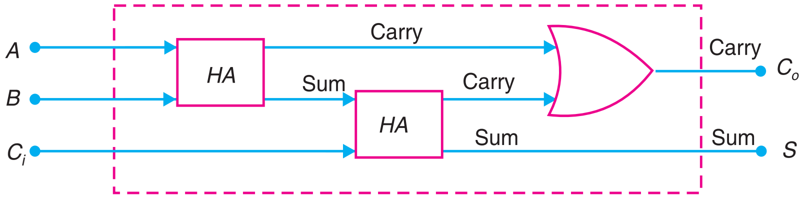


Fig. 71.2

It can add three bits at a time.

Fig. 71.3 shows the symbol of a full-adder. It has two inputs *A* and *B*, plus a third input *C*. Input *C* is also called the CARRY IN (C_i). It comes from a lower-order column. There are two outputs, *Sum* and *Carry*. The output *carry* is also called the *CARRY OUT* (C_o).

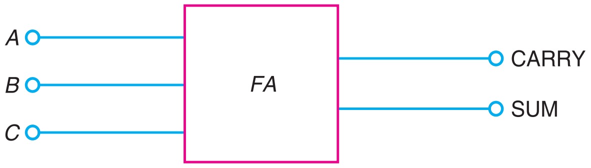


Fig. 71.3

It goes to the next higher column.

The truth-table of a *FA* is given in Table 2.

It has three inputs and two outputs.

TRUTH TABLE 2

Input			*Output*	
A	*B*	C_i	*Sum S*	*Carry* C_o
0	0	0	0	0
0	0	1	1	0
0	1	0	1	0
0	1	1	0	1
1	0	0	1	0
1	0	1	0	1
1	1	0	0	1
1	1	1	1	1

Here C_i represents a carry from a preceding stage.

C_o represents the carry to be added to the next stage.

Working

To illustrate its operation, let us take two examples:

1. $A = 1, B = 1, C_i = 0$

Fig. 71.4 shows the full adder with these three inputs.

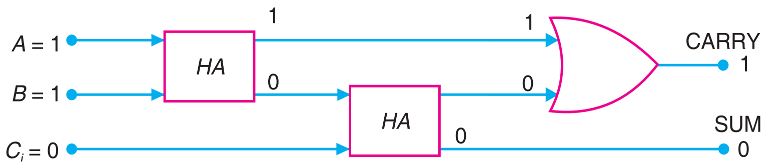

A = 1
B = 1
C_i = 0
HA
HA
1
0
1
0
0
CARRY
1
SUM
0

Fig. 71.4

(*i*) The output of first half adder consists of a sum of 0 with a carry of 1.

(*ii*) The sum 0 of first half adder and carry 0 when fed to the second half adder, give a sum of 0 with a carry of 0.

(*iii*) The carry outputs of both the half adders is fed into the input of OR gate.

The final output is : SUM 0, CARRY 1.

We get the same result from binary addition: $1 + 1 + 0 = 10_2$.

2. $A = 1, B = 1, C_i = 1.$

(*i*) The output of the first half adder is a sum of 0 with a carry 1 (Fig. 71.5).

(*ii*) The output of second half adder is a sum of 1 with a carry of 0.

(*iii*) The final output is SUM 1 with a CARRY 1.

We get the same result from binary addition: $1 + 1 + 1 = 11_2$.

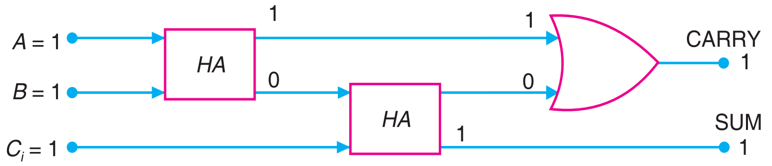

A = 1
B = 1
C_i = 1
HA
HA
1
0
1
0
1
CARRY
1
SUM
1

Fig. 71.5

Fig. 71.6 shows the actual circuit of a full adder.

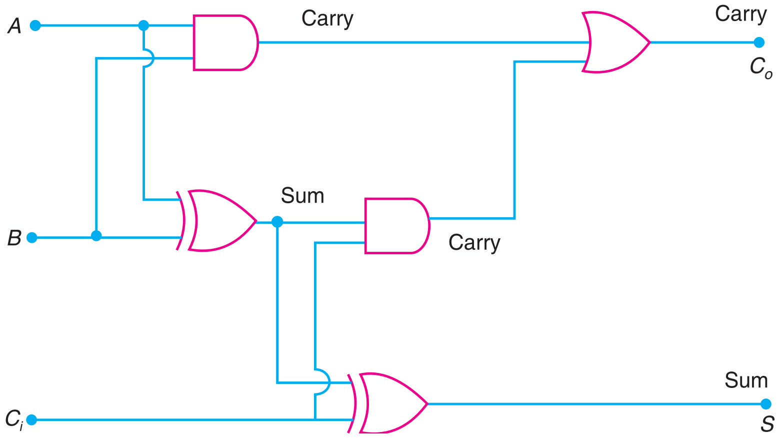

A
B
C_i
Carry
Sum
Carry
Carry
C_o
Sum
S

Fig. 71.6

The final sum is given by the XOR gate of the second half adder.

The final carry is given by the OR gate.

71.3 Parallel Binary Adder

A parallel adder is used to add two numbers in parallel form and to produce the sum bits as parallel outputs. By cascading a number of full adders, we can get an *n*-bit binary adder circuit which is called a *parallel adder.*

Suppose we are required to add 4-bit binary numbers $A_4 A_3 A_2 A_1$ and $B_4 B_3 B_2 B_1$. Then we get a sum $S_4 S_3 S_2 S_1$ such that

$$\begin{array}{r} A_4 A_3 A_2 A_1 \\ +\ B_4 B_3 B_2 B_1 \\ \hline S_5 S_4 S_3 S_2 S_1 \end{array}$$

Here, S_5 indicates overflow carry if the sum exceeds four bits. For adding them (two 4-bit numbers), we need *four full adders* (FA) *connected in parallel*. The first adder could, however, be a half adder (HA) because only two bits are to be added and there is no carry. But all subsequent columns need full adder because we have to handle three bits at a time (two binary digits and a carry generated by the previous column).

Fig. 71.7 shows a parallel four-bit binary adder.

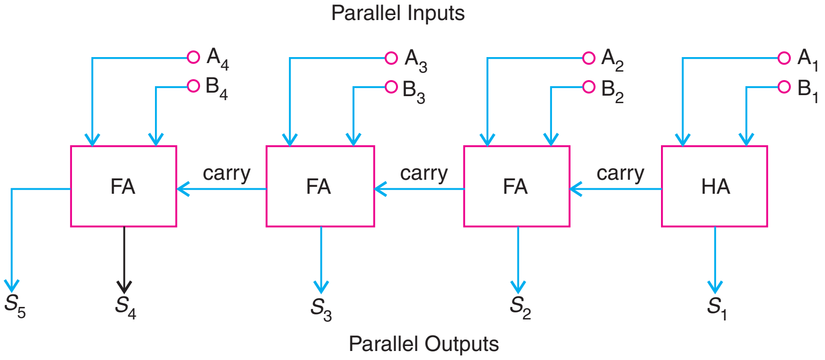


Fig. 71.7

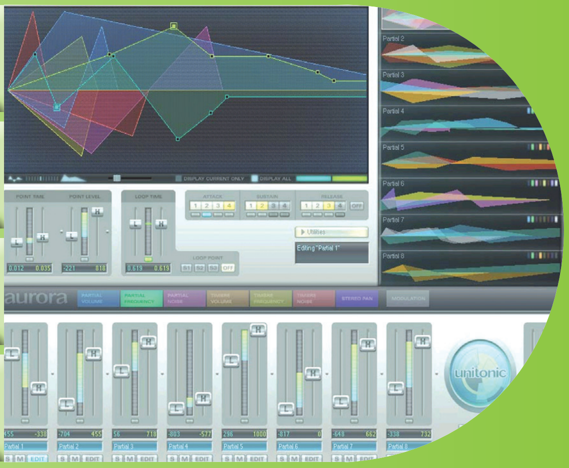
To increase the capacity, more full adders may be connected. For example, for adding 6-bit numbers, we will have to add two more full adders to the left, thus making a total of six.

EXERCISE

1.	Explain the working of a half adder with an example.	(*Andhra,* 94)
2.	Draw the logic diagram of a half adder and write its truth table.	(*Osmania,* 93)
3.	Explain the working of a full adder with the help of a circuit diagram.	(*Andhra,* 93, 94)

CHAPTER

72



BJT TRANSISTOR MODELING

AT A GLANCE

72.1	Introduction	**72.2**	The Important Parameters: Z_i, Z_o, A_v, A_i
72.3	The r_e Transistor Model		

72.1 Introduction

The key to transistor small-signal analysis is the use of equivalent circuits (models).

A model is the combination of circuit elements, properly chosen, that best approximates the actual behavior of a semiconductor device under specific operating conditions.

There are two models commonly used in the small-signal ac analysis of transistor networks : the r_e model and the *hybrid equivalent* model.

This chapter introduces the r_e model.

The ac equivalent of a network is obtained by :

1. *Setting all dc sources to zero and replacing them by a short-circuit equivalent.*
2. *Replacing all capacitors by a short-circuit equivalent.*
3. *Removing all elements bypassed by the short-circuit equivalents introduced by steps 1 and 2.*
4. *Redrawing the network in a more convenient and logical form.*

72.2 The Important Parameters: Z_i, Z_o, A_v, A_i

The BJT acts as a two-port (two pairs of terminals) system (Fig. 72.1). The input side (the side to which the signal is normally applied) is to the left. The output side (where the load is connected) is to the right. For most electrical and electronic systems the general flow is usually from the left to the right. For both sets of terminals the impedance between each pair of terminals under normal operating conditions is quite important.

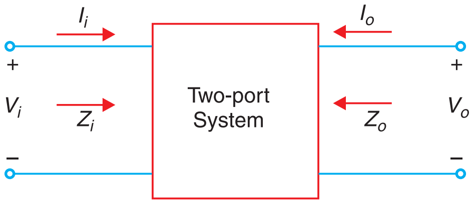


Fig. 72.1

(*i*) Input Impedance, Z_i

For the input side, the input impedance Z_i is defined by Ohm's law as the following :

$$Z_i = \frac{V_i}{I_i} \qquad \text{...(1)}$$

If the input signal V_i is changed, the current I_i can be computed using the same level of input impedance. In other words :

For small-signal analysis, once the input impedance has been determined the same numerical value can be used for changing levels of applied signal.

In particular, for frequencies in the low to mid-range (typically ≤ 100 *kHz*):

The input impedance of a BJT transistor amplifier is purely resistive in nature, and depending on the manner in which the transistor is employed, can vary from a few ohms to megohms.

In addition :

An ohmmeter cannot be used to measure the small-signal ac input impedance since the ohmmeter operates in the dc mode.

Determination of Z_i

Eq. (1) provides a method for measuring the input resistance in the ac domain. In Fig. 72.2 a sensing resistor has been added to the input side to permit a determination of I_i using Ohm's law. An oscilloscope or sensitive digital multimeter (DMM) can be used to measure the voltage V_s and V_i. Both voltages can be the peak-to-peak, peak, or rms values, as long as both levels use the same standard. The input impedance is then determined in the following manner :

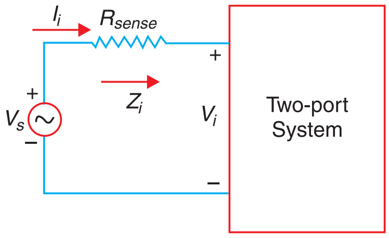


Fig. 72.2

$$I_i = \frac{V_s - V_i}{R_{sense}} \qquad \text{...(2)}$$

and

$$Z_i = \frac{V_i}{I_i} \qquad \text{...(3)}$$

(*ii*) Output Impedance, Z_o

The output impedance of any amplifier is defined as the ratio of output voltage to the output current keeping the input current zero (input terminals left open).

Determination of Z_o

The output impedance is determined at the output terminals looking back into the system with the applied signal set to zero.

In Fig. 72.3, the applied signal has been set to zero volts.

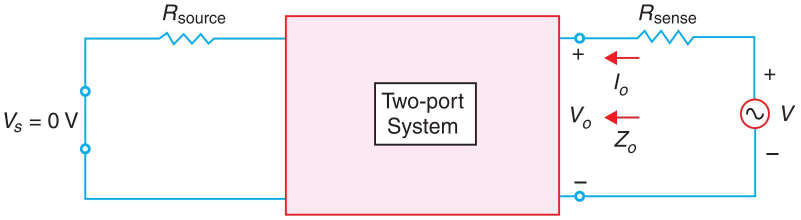


Fig. 72.3

A signal, V_s, is applied to the output terminals. The level of V_o is measured with an oscilloscope or sensitive DMM. The output impedance is then determined in the following manner :

$$I_o = \frac{V - V_o}{R_{sense}} \qquad \text{...(4)}$$

and $$Z_o = \frac{V_o}{I_o} \qquad \text{...(5)}$$

In particular for frequencies in the low to mid-range (typically ≤ 100 kHz) :

The output impedance of a BJT transistor amplifier is resistive in nature and depending on the configuration and the placement of the resistive elements.

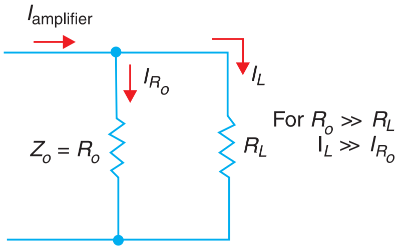


Fig. 72.4

Z_o, can vary from a few ohms to a level that can exceed 2 MΩ.

In addition :

An ohmmeter cannot be used to measure the small-signal ac output impedance since the ohmmeter operates in the dc mode.

Effect of $Z_o = R_o$ on the load or output current I_L.

For amplifier configurations where significant gain in current is desired, the level of Z_o should be as large as possible. If $Z_o \gg R_L$, the majority of the amplifier output current will pass on to the load (Fig. 72.4). Z_o is frequently so large compared to R_L that it can be replaced by an open-circuit equivalent.

(*iii*) Voltage Gain, A_v

The voltage gain of any amplifier is defined as the ratio of the output voltage to the input voltage.

The small-signal ac voltage gain of an amplifier is

$$A_v = \frac{V_o}{V_i} \qquad \text{...(6)}$$

Determination of the no-load voltage gain.

For the system of Fig. 72.5, a load has not been connected to the output terminals. The level of gain determined by Eq. (6) is referred to as the no-load voltage gain. That is,

$$A_{v_{NL}} = \left.\frac{V_o}{V_i}\right|_{R_L = \infty\Omega(\text{open}-\textit{circuit})} \qquad \text{...(7)}$$

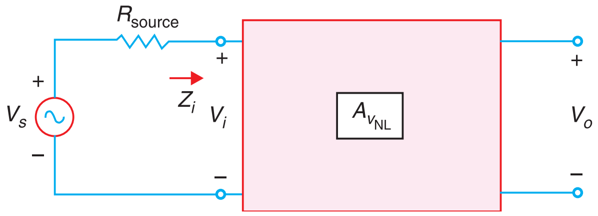


Fig. 72.5

For transistor amplifiers, the no-load voltage gain is greater than the loaded voltage gain.

Calculation of the gain V_o/V_s

For the system of Fig. 72.5, having a source resistance R_s, the level of V_i would first have to be determined using the voltage-divider rule.

$$V_i = \frac{Z_i V_s}{Z_i + R_s} \quad \text{or} \quad \frac{V_i}{V_s} = \frac{Z_i}{Z_i + R_s}$$

$$A_{v_s} = \frac{V_o}{V_s} = \frac{V_i}{V_s} \cdot \frac{V_o}{V_i}$$

$$\therefore \quad A_{v_s} = \frac{V_o}{V_s} = \frac{Z_i}{Z_i + R_s} A_{v_{\mathrm{NL}}} \qquad ...(8)$$

Experimentally, the voltage gain A_{v_s} or $A_{v_{\mathrm{NL}}}$ can be determined simply by measuring the appropriate voltage levels with an oscilloscope or sensitive DMM and substituting into the appropriate equation.

Depending on the configuration, the magnitude of the voltage gain for a loaded single-stage transistor amplifier typically ranges from just less than 1 to a few hundred. A multistage (multiunit) system, however, can have a voltage gain in the thousands.

(*iv*) Current Gain, A_i

The current gain is defined by

$$A_i = \frac{I_o}{I_i} \qquad ...(9)$$

For BJT amplifiers, the current gain typically ranges from a level just less than 1 to a level that may exceed 100.

Determination of the loaded current gain.

For the loaded situation of Fig. 72.6,

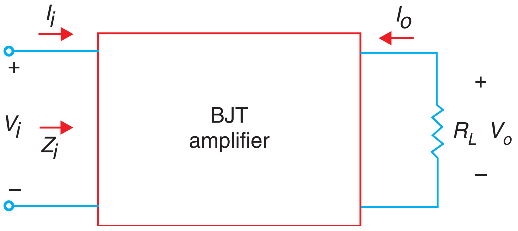


Fig. 72.6

$$I_i = \frac{V_i}{Z_i} \quad \text{and} \quad I_o = -\frac{V_o}{R_L}$$

$$A_i = \frac{I_o}{I_i} = -\frac{V_o / R_L}{V_i / Z_i} = -\frac{V_o Z_i}{V_i R_L}$$

$$\therefore \quad A_i = -A_v \frac{Z_i}{R_L} \qquad ...(10)$$

Eq. (10) allows the determination of the current gain from the voltage gain and the impedance levels.

Phase Relationship

For the typical transistor amplifier at frequencies that permit ignoring the effects of the reactive elements, the input and output signals are either 180° out of phase or in phase.

72.3 The r_e Transistor Model

The r_e model employs a diode and controlled current source to duplicate the behavior of a transistor in the region of interest. A current-controlled current source is one where the parameters of the current source are controlled by a current elsewhere in the network. In general:

BJT transistor amplifiers are referred to as current-controlled devices.

72.3.1. Common Base Configuration

In Fig. 72.7 (*a*), a common-base *pnp* transistor has been inserted within the two-port structure. In Fig. 72.7 (*b*), the r_e model for the transistor has been placed between the same four terminals.

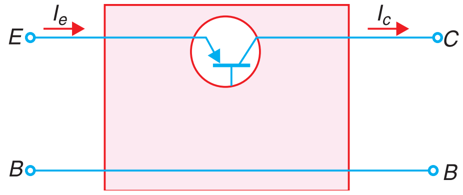


Fig. 72.7 (a)

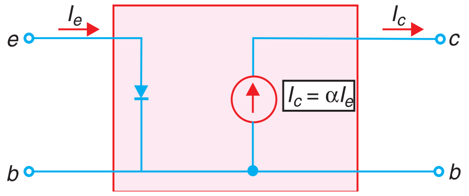


Fig. 72.7 (b)

One junction of an operating transistor is forward-biased while the other is reverse-biased.

(*i*) The forward-biased junction will behave much like a diode (ignoring the effects of changing levels of V_{CE}). For the base-to-emitter junction of the transistor of Fig. 72.7 (*a*), the diode equivalence of Fig. 72.7 (*b*) between the same two terminals is quite appropriate.

(*ii*) The output characteristics for a common-base transistor amplifier indicate that $I_c \cong I_e$ (as derived from $I_c = \alpha I_e$) for the range of values of V_{CE}. The current source of Fig. 72.7(*b*) establishes the fact that $I_c = \alpha I_e$ with the controlling current I_e appearing in the input side of the equivalent circuit as dictated by Fig. 72.7 (*a*). We have therefore established an equivalence at the input and output terminals with the current-controlled source, providing a link between the two. So the model of Fig. 72.7 (*b*) is a valid model of the actual device.

The ac resistance of a diode can be determined by the equation $r_{ac} = 26\ \text{m}V/I_D$, where I_D is the dc current through the diode at the Q (quiescent) point. This same equation can be used to find the ac resistance of the diode of Fig. 72.7 (*b*) if we simply substitute the emitter current as follows:

$$r_e = \frac{26\ mV}{I_E} \qquad \text{...(1)}$$

The subscript *e* of r_e emphasizes that it is the dc level of emitter current that determines the ac level of the resistance of the diode. Substituting the resulting value of r_e in Fig. 72.7 (*b*) results in the model of Fig. 72.8.

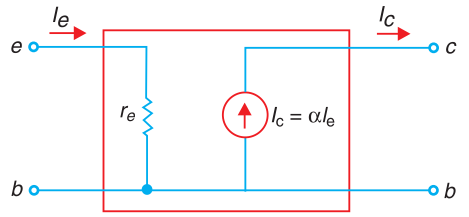


Fig. 72.8

Fig. 72.8 represents common-base r_e equivalent circuit.

Determination of Z_i, Z_o, A_v, A_i for the r_e Transistor Model

(*i*) Due to the isolation that exists between input and output circuits of Fig. 72.8, the input impedance Z_i for the common-base configuration of a transistor is simply r_e. That is,

$$\boxed{Z_i = r_e}_{CB} \qquad \text{...(2)}$$

For the common-base configuration, typical values of Z_i range from a few ohms to a maximum of about 50 Ω.

(*ii*) Consider the r_e Transistor Model of Fig. 72.8. If we set the signal to zero, then $I_e = 0\ A$ and $I_c = \alpha I_e = \alpha(0A) = 0A$, resulting in an open-circuit equivalence at the output terminals.

$$\therefore \qquad \boxed{Z_o \approx \infty\ \Omega}_{CB} \qquad \text{...(3)}$$

For the common-base configuration, typical values of Z_o are in the megohm range.

Defining Z_o. The output resistance of the *C–B* configuration is determined by the slope of the characteristic lines of the output characteristics (Fig. 72.9). If Z_o is measured graphically or experimentally, levels typically in the range 1- to 2-MΩ are obtained. If we assume the lines to be perfectly horizontal (an excellent approximation), $Z_o \cong \infty\ \Omega$.

In general, for the common-base configuration the input impedance is relatively small and the output impedance quite high.

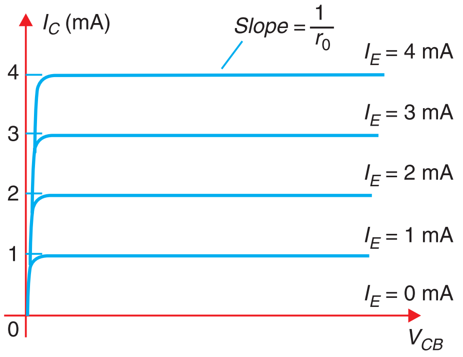


Fig. 72.9

(*iii*) We shall determine the voltage gain $A_v = V_o/V_i$ for the network of Fig. 72.10.

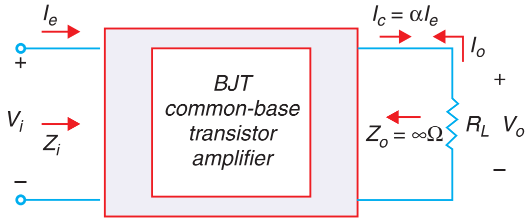


Fig. 72.10

$$V_o = -I_oR_L = -(-I_c)R_L = \alpha I_eR_L$$

and
$$V_i = I_eZ_i = I_er_e$$

$$\therefore \qquad A_v = \frac{V_o}{V_i} = \frac{\alpha I_eR_L}{I_er_e}$$

and
$$\boxed{A_v = \frac{\alpha R_L}{r_e} \cong \frac{R_L}{r_e}}_{CB} \qquad \text{...(4)}$$

(*iv*) Current gain,
$$A_i = \frac{I_o}{I_i} = \frac{-I_c}{I_e} = -\frac{\alpha I_e}{I_e}$$

and
$$\boxed{A_i = -\alpha \approx -1}_{CB} \qquad \text{...(5)}$$

The fact that the polarity of the voltage V_o as determined by the current I_c is the same as defined by Fig. 72.10 (*i.e.,* the negative side is at ground potential) reveals that V_o and V_i are *in phase* for the common-base configuration.

Fig. 72.11 shows the approximate *model for a common-base npn transistor configuration.*

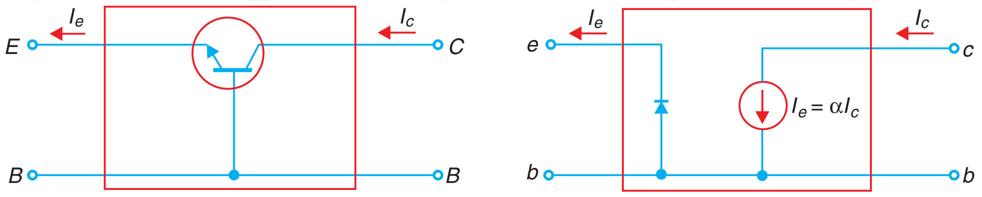


Fig. 72.11

72.3.2. Common Emitter Configuration

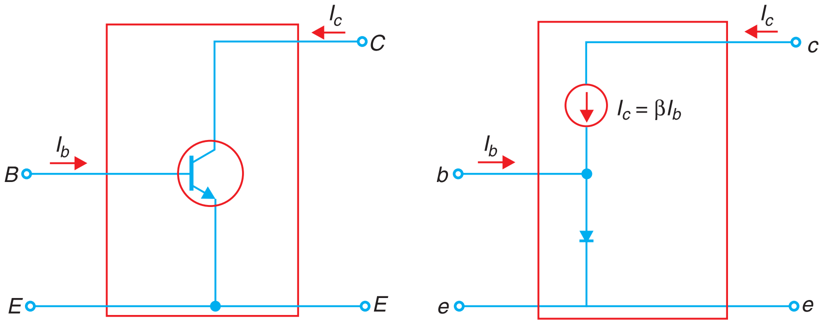
Fig. 72.12 (*a*) shows the *C–E Configuration for an npn* transistor.

(*i*) The input terminals are the base and emitter terminals.

(*ii*) The output terminals are the collector and emitter terminals.

The emitter terminal is common between the input and output ports of the amplifier.

Fig. 72.12 (*b*) shows the approximate model for the configuration of Fig. 72.12 (*a*).



(a) (b)

Fig. 72.12

The controlled-current source is connected between the collector and base terminals.

The diode is connected between the base and emitter terminals.

The base current I_b is the input current. The output current is I_c.

The base and collector currents are related by the following equation:

$$I_c = \beta I_b \qquad ...(1)$$

The current through the diode is therefore determined by

$$I_e = I_c + I_b = \beta I_b + I_b$$

$$I_e = (\beta + 1) I_b \qquad ...(2)$$

Since ac $\beta >> 1$, we will use the following approximation for the current analysis:

$$I_e \cong \beta I_b \qquad ...(3)$$

(*i*) Determining Z_i using the approximate model. The input impedance is determined by the following ratio:

$$Z_i = \frac{V_i}{I_i} = \frac{V_{be}}{I_b}$$

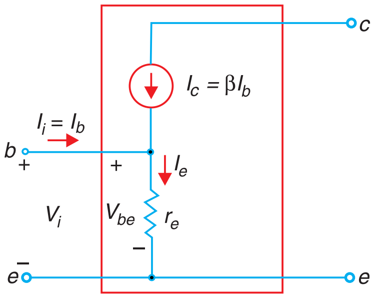


Fig. 72.13

The voltage V_{be} is across the diode resistance (Fig. 72.13). The level of r_e is determined by the dc current I_E.

Using Ohm's law gives

$$V_i = V_{be} = I_e r_e \cong \beta I_b r_e$$

$$\therefore \qquad Z_i = \frac{V_{be}}{I_b} \approx \frac{\beta I_b r_e}{I_b}$$

and $$\boxed{Z_i \simeq \beta r_e}_{CE} \qquad ...(4)$$

Impact of r_e on input impedance

Input impedance = β times the value of r_e. That is, a resistive element in the emitter leg is reflected into the input circuit by a multiplying factor β (Fig. 72.14).

EXAMPLE. *If $r_e = 6.5\ \Omega$ and $\beta = 160$, the input impedance has increased to a level of*

$Z_i = \beta r_e = (160)\ (6.5\ \Omega) =$ **1.04 $k\Omega$**

For the C–E configuration, typical values of Z_i defined by βr_e range from a few hundred ohms to the kilohm range, with maximums of about 6 – 7 $k\Omega$.

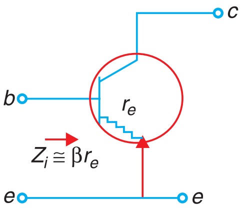


Fig. 72.14

(*ii*) Defining r_o for the C–E configuration

Fig. 72.15 shows the output characteristics of a silicon transistor. The slope of the curves increases with increase in collector current. The steeper the slope, the less the level of output impedance (Z_o).

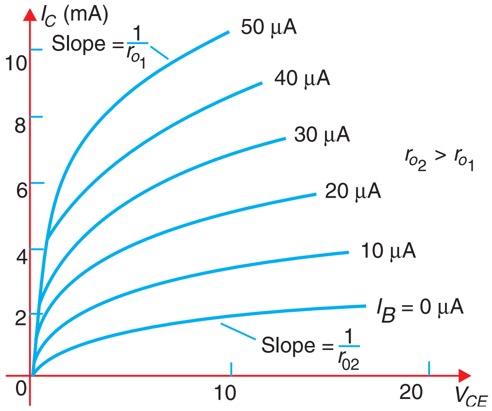


Fig. 72.15

Including r_0 in the transistor equivalent circuit

The r_e model of Fig. 72.12 does not include an output impedance, but if available from a graphical analysis or from data sheets it can be included as shown in Fig. 72.16.

For the C–E configuration, typical values of Z_o are in the range of 40 to 50$k\Omega$.

For the model of Fig. 72.16, if the applied signal is set to zero, the current I_c is 0*A*.

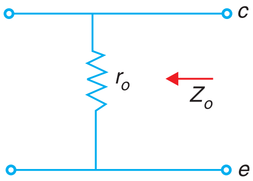


Fig. 72.16

The output impedance is

$$\boxed{Z_o = r_o}_{CE} \qquad ...(5)$$

If the contribution due to r_o is ignored as in the r_e model, the output impedance is defined by $Z_o = \infty\ \Omega$.

(*iii*) Determination of voltage gain for the C–E transistor amplifier.

We assume that $Z_o = \infty\Omega$ (Fig. 72.17).

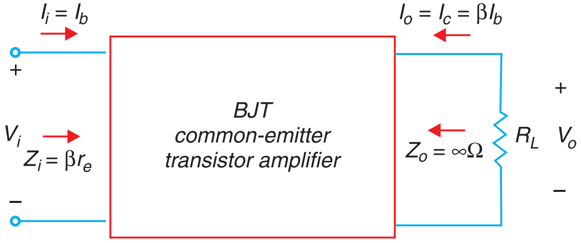


Fig. 72.17

For the defined direction of I_o and polarity of V_o,

$$V_o = -I_o R_L$$

The minus sign simply reflects the fact that the direction of I_o in Fig. 72.17 would establish a voltage V_o with the opposite polarity.

$$V_o = -I_o R_L = -I_c R_L = -\beta I_b R_L$$

and
$$V_i = I_i Z_i = I_b \beta r_e$$

$$A_v = \frac{V_o}{V_i} = -\frac{\beta I_b R_L}{I_b \beta r_e}$$

$$\therefore \qquad \boxed{A_v = -\frac{R_L}{r_e}}_{CE,\, r_o = \infty \Omega} \qquad \text{...(6)}$$

The minus sign for the voltage gain reveals that the output and input voltages are 180° out of phase.

(*iv*) Determination of current gain

The current gain for the configuration of Fig. 72.17 is

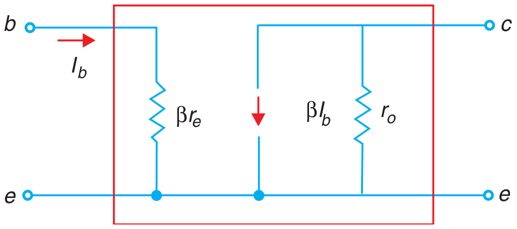


Fig. 72.18

$$A_i = \frac{I_o}{I_i} = \frac{I_c}{I_b} = \frac{\beta I_b}{I_b}$$

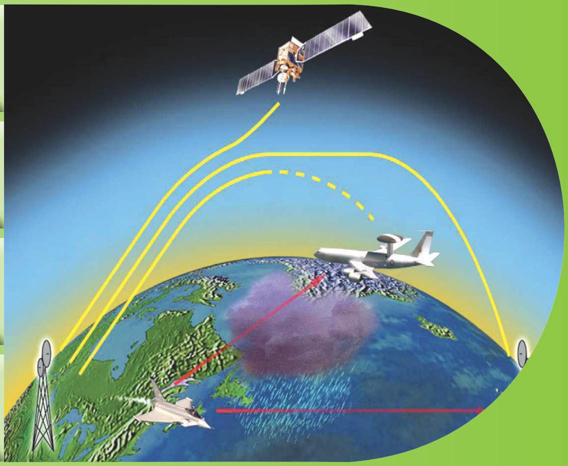
$$\therefore \qquad \boxed{A_i = \beta}_{CE,\, r_o = \infty\, \Omega} \qquad \text{...(7)}$$

Using the facts that the input impedance is βr_e, the collector current is βI_b, and the output impedance is r_o, Fig. 72.18 shows the r_e model for the *common-emitter transistor configuration.*

The *C–E* configuration has a moderate level of input impedance, a high voltage and current gain, and an output impedance that may have to be included in the network analysis.

Common Collector Configuration

For the *C–C* configuration the model defined for the *C–E* configuration of Fig. 72.12 is normally applied.



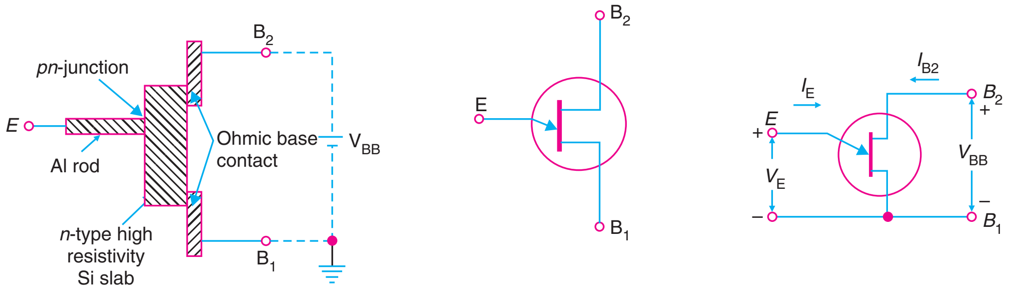
CHAPTER 73

UNIJUNCTION TRANSISTOR (UJT)

AT A GLANCE			
73.1	Construction	**73.2**	Equivalent Circuit of a UJT
73.3	Operation of UJT	**73.4**	Static Emitter–Characteristic Curve for a UJT
73.5	Applications of UJT		

73.1 Construction

Fig. 73.1 (*a*) shows the basic construction of a unijunction transistor (UJT).



(*a*) Basic construction of a UJT (*b*) Symbolic representation (*c*) Biasing arrangement

Fig. 73.1

It is a three-terminal device. A slab of lightly doped (increased resistance characteristic) *n*-type silicon material has two base contacts attached to both ends of one surface and an aluminium rod alloyed to the opposite surface. The *p*-*n* junction of the device is formed at the boundary of the aluminium rod and the *n*-type silicon slab. The single *p*-*n* junction accounts for the terminology *unijunction*. The aluminium rod is alloyed to the silicon slab at a point closer to the base 2 contact than the base 1 contact. Further, base 2 terminal is made positive with respect to the base 1 terminal by V_{BB} volts.

Fig. 73.1 (*b*) shows the circuit symbol of a UJT. Note that the emitter leg is drawn at an angle to the vertical line representing the slab of *n*-type material. The arrowhead is pointing in the direction of conventional current (hole) flow when the device is in the forward-biased, active, or conducting state.

The basic biasing arrangement for the unijunction transistor is provided in Fig. 73.1 (*c*).

73.2 Equivalent Circuit of a UJT

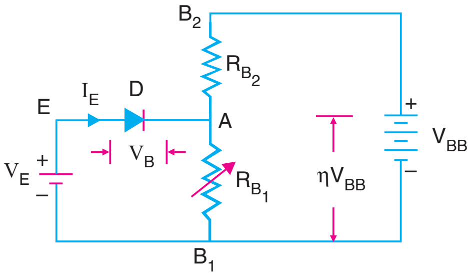


Fig. 73.2

The circuit equivalent of the UJT is shown in Fig. 73.2. Note the relative simplicity of this equivalent circuit: two resistors (one fixed, one variable) and a single diode.

The resistance of the silicon bar is represented by two series resistors: R_{B_2} is the resistance of the base-2 portion.

The resistance R_{B_1} is shown as a variable resistor since its magnitude will vary with the current I_E. For a representative UJT, R_{B_1} may vary from 5 kΩ down to 50 Ω for a corresponding change of I_E from 0 to 50 μ A.

The interbase resistance R_{BB} is the resistance between terminals B_1 and B_2 when $I_E = 0$. In equation form,

$$R_{BB} = (R_{B_1} + R_{B_2})|_{I_E} = 0 \qquad \text{...(1)}$$

Typically, R_{BB} ranges from 4 to 10 kΩ.

The position of the aluminium rod will determine the relative values of R_{B_1} and R_{B_2} with $I_E = 0$.

A battery V_{BB} is connected across B_2 B_1 of equivalent circuit of UJT. Part of V_{BB} is dropped over R_{B2} and part on R_{B1}.

The magnitude of $V_{R_{B_1}}$ (with $I_E = 0$ *i.e.*, emitter open) is determined by the voltage-divider rule.

$$V_{R_{B_1}} = \frac{R_{B_1} V_{BB}}{R_{B_1} + R_{B_2}} = \eta V_{BB}\Bigg|_{I_E = 0} \qquad \text{...(2)}$$

Here, η is called the *intrinsic stand-off ratio* of the device. It is defined by

$$\eta = \frac{R_{B_1}}{R_{B_1} + R_{B_2}}\Bigg|_{I_E = 0} = \frac{R_{B_1}}{R_{BB}} \qquad \text{...(3)}$$

For applied emitter potentials (V_E) greater than $V_{R_{B_1}} (= \eta V_{BB})$ by the forward voltage drop of the diode V_D $(0.35 \rightarrow 0.70\ V)$, the diode will fire. Assume the short-circuit representation (on an ideal basis), and I_E will begin to flow through R_{B1}. In equation form, the emitter firing potential is given by

$$V_P = \eta\, V_{BB} + V_D \qquad \text{...(4)}$$

73.3 Operation of UJT

Fig. 73.3 shows the basic circuit operation of UJT.

Normally the base 1 (B_1) is grounded. A positive voltage V_{BB} is applied at the base 2 (B_2). When there is no emitter current, this voltage V_{BB} produces a uniform drop across the 5000 to 10,000 ohms internal resistance of silicon bar. The resultant current through the bar will produce a voltage drop

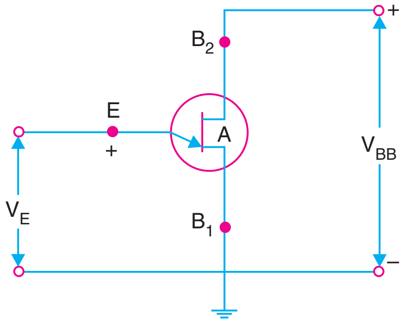


Fig. 73.3

η V_{BB} between points A and B_1 of the bar where A denotes the point, where the P region is formed. So long as the voltage V_E applied to the emitter with respect to B_1 is less than V_A or V_P (= η V_{BB}), the *p-n* junction remains reverse biased and only a very small reverse current I_E flows in the emitter circuit. However, at the instant the emitter voltage V_E becomes greater than ηV_{BB}, the junction becomes forward biased and the emitter current I_E increases. The increase is quite steep because this current is due to the holes injected into the silicon, which move from the emitter to base 1. The presence of these holes in the silicon bar attracts a considerable number of electrons to this region and, in consequence, causes the bar resistance between emitter and base 1 to drop sharply. The upshot is that an increase in emitter current causes a decrease in the resistance and in the emitter voltage, which is known as a negative resistance characteristic.

73.4 Static Emitter – Characteristic Curve for a UJT

Fig. 73.4 shows the characteristics of a representative unijunction transistor for $V_{BB} = 10\ V$.

If V_{BB} is made zero, and a voltage is applied to V_E, the resulting current I_E that flows gives the emitter to base-1 diode characteristic ($I_{B2} = 0$).

Note that for emitter potentials to the left of the peak point, the magnitude of I_E is never greater than I_{E0} (measured in microamperes). The current I_{E0} corresponds very closely with the reverse leakage current I_{C0} of the conventional bipolar transistor. This region, is called the *cut off region*. Conduction is established at $V_E = V_P$. After V_P, an attempt to increase V_E is followed by a sudden increase in emitter current I_E with a corresponding decrease in V_E. This corresponds to a *negative resistance* region. After the negative resistance region, the valley point is reached. Further increase in I_E places the device in the *saturation region* (with characteristic approaching that of a semiconductor diode).

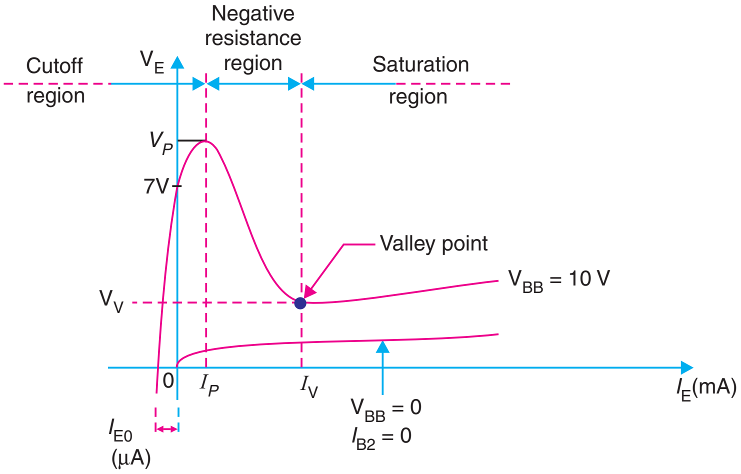


Fig. 73.4

The decrease in resistance in the active region is due to the holes injected into the *n*-type slab from the aluminium *p*-type rod when conduction is established. The increased hole content in the *n*-type material will result in an increase in the number of free electrons in the slab, producing an increase in conductivity (G) and a corresponding drop in resistance ($R \downarrow = 1/G \uparrow$). Three other important parameters for the UJT are I_P, V_V and I_V. Each is indicated on Fig. 73.4.

Fig. 73.5 shows the typical family of *static emitter characteristic curves* of a UJT for various V_{BB} (voltage between the bases) values.

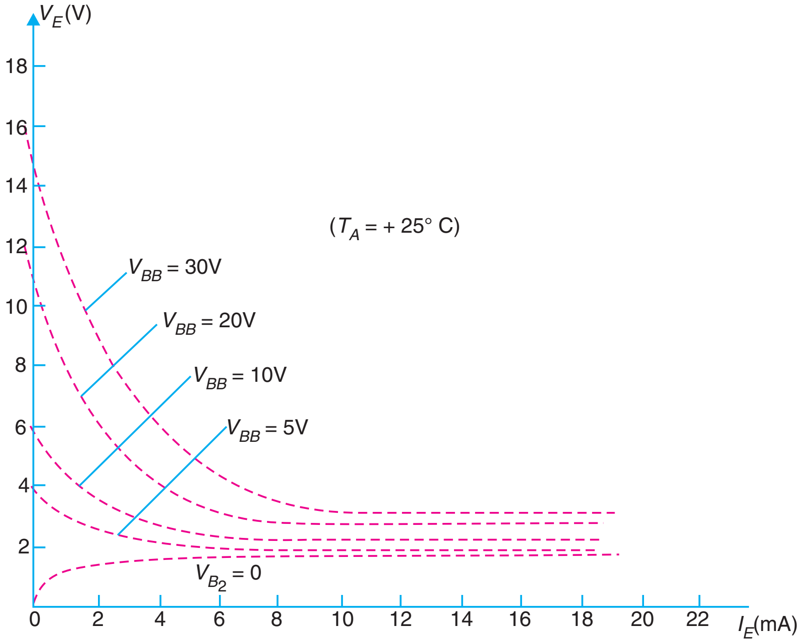


Fig. 73.5

Note that I_{E0} (μA) is not in evidence since the horizontal scale is in milliamperes.

The intersection of each curve with the vertical axis is the corresponding value of V_P. For fixed values of η and V_D, the magnitude of V_P will vary as V_{BB}, that is,

$$V_P \uparrow = \underbrace{\eta V_{BB} \uparrow + V_D}_{\text{fixed}}$$

73.5 Applications of UJT

(1) UJT relaxation oscillator

The UJT can be used in a single stage oscillator circuit to provide a pulse signal suitable for digital-circuit applications.

Circuit Details

Fig. 73.6 shows the circuit of the UJT relaxation oscillator.

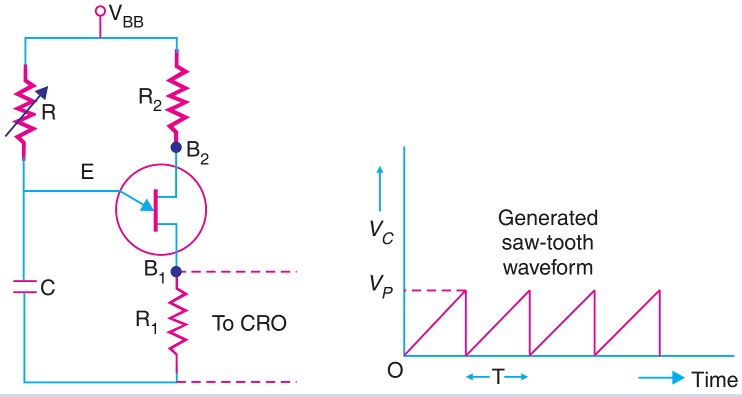

V_{BB}
R
R_2
B_2
E
C
B_1
R_1
To CRO
V_C
V_P
Generated saw-tooth waveform
O
T
Time

Fig. 73.6

It consists of a UJT and a capacitor C which is charged through a variable resistance R when V_{BB} is switched on. R_1 and R_2 are the external resistances such that $R_1 << R_2$. R_1 is less than 50 Ω. The inclusion of these resistances provide spike waveforms. Here the negative resistance region of the UJT is used. The discharge of the capacitor through UJT develops a saw-tooth output. Resistor R and capacitor C are the timing components that set the circuit oscillating rate.

Circuit operation

When the circuit is switched on, the capacitor C starts charging through R. During charging period, the voltage across capacitor increases in an exponential relation till it reaches the peak point voltage V_P. At this instant of time, UJT switches to its low resistance conducting mode. So, the capacitor C discharges suddenly through UJT and R_1. Discharge time constant is very small due to small value of R_1. So C discharges abruptly. As the capacitor voltage moves back to zero, the emitter ceases to conduct and UJT is switched off. The capacitor starts charging again and the next cycle begins. The output saw-tooth waveform frequency can be varied by changing the value of R.

Frequency of oscillation

The frequency of saw-tooth wave can be calculated as follows :

The voltage across capacitor during charging at any instant t is given by

$$v_c = V_{BB}(1 - e^{-t/\tau})$$

where τ is charging time constant = $R\,C$.

This voltage is applied to the emitter. So

$$v_c = V_E = V_{BB}(1 - e^{-t/\tau})$$

When capacitor charges to peak voltage V_P, UJT triggers. The UJT triggers when $V_P = \eta\, V_{BB}$.

If the capacitor takes a time T to charge to firing potential V_P, then

$$V_P = V_{BB}(1 - e^{-T/CR})$$

or
$$\frac{V_P}{V_{BB}} = \eta = (1 - e^{-T/CR})$$

or
$$e^{-T/CR} = 1 - \eta$$

or
$$T = CR \log_e\left(\frac{1}{1-\eta}\right)$$

$$\text{Frequency of oscillation, } f = \frac{1}{T} = \frac{1}{RC\log_e\left(\frac{1}{1-\eta}\right)}.$$

The oscillating frequency may be calculated using the above equation, which includes the UJT intrinsic stand-off ratio η as a factor (in addition to R and C) in the oscillator operating frequency.

(2) UJT as over voltage detector

Fig. 73.7 shows a simple d.c. over voltage detector. A warning lamp L is connected between emitter E and B_1 circuit.

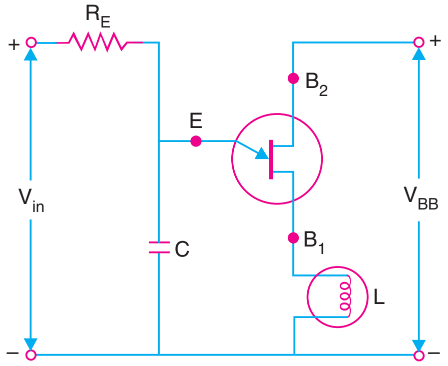


Fig. 73.7

Principle

As long as the input voltage is less than peak-point voltage V_P of UJT, the device remains switched off. When the input voltage exceeds V_P, the device is switched on.

Working

Suppose we want to keep certain voltage constant say at value V_i. Now we choose V_{BB} in such a way that this is slightly less than V_i. In case V_i exceeds V_P, then UJT is switched on. As a result, the capacitor starts discharging through low resistance path between E and B_1. The current flowing in pilot lamp L lights it. This gives an indication of over voltage in the circuit.

CHAPTER 74

FEEDBACK AMPLIFIERS

AT A GLANCE

74.1 Feedback Principle
74.2 The Effect of Feedback on the Gain of an Amplifier
74.3 Advantages of Negative Feedback
74.4 Practical Feedback Circuits

74.1 Feedback Principle

In an amplifier, a small part of the output signal may be fed to the input circuit by suitable means. This is known as feedback. Feedback is a process in which there is a transfer of energy from output of an electronic system to its input. The amount of feedback is determined by the feedback network.

(*a*) Positive feedback amplifier : When the feedback energy (Voltage or current) is in phase with the input signal and thus aids it, it is called *positive feedback*. When the input and feedback signals are in same phase, the magnitude of the input signal enhances as an effect of the feedback signal. The amplifier based on positive feedback is called *positive feedback* amplifier. They are generally used as oscillators.

(*b*) Negative feedback amplifier : When the feedback energy (voltage or current) is out of phase with the input signal and thus opposes it, it is called *negative feedback.* When the input and feedback signals are completely out of phase, the magnitude of the input signal diminishes by the superimposition of feedback signal. The amplifier based on negative feedback is called negative feedback amplifier.

The feedback amplifiers are classified into two general classes : voltage feedback and current feedback. In voltage feedback, the voltage fed back to the input terminals is proportional to the output voltage. In current feedback, the voltage fed back to the input terminals is proportional to the current through the load.

The amount of amplifier gain with feedback is called the *closed loop gain*. Without feedback, the gain is called *open loop gain*.

74.2 The Effect of Feedback on the Gain of an Amplifier

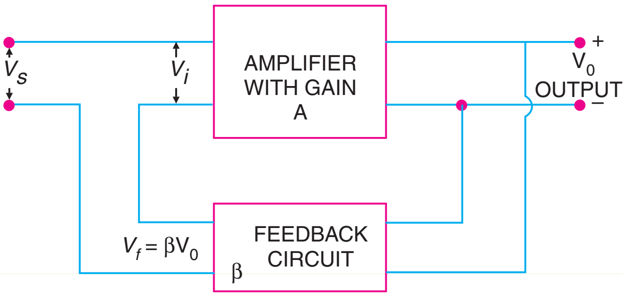


Fig. 74.1

Fig. 74.1 shows the block diagram of an amplifier with feedback loop. The feedback circuit usually consists of resistors. A fraction β of the output voltage is fed back into the input circuit through the feedback circuit. Let A be the gain of the amplifier *i.e.*, the ratio of output voltage V_0 to the input voltage V_i.

$$A = \frac{V_0}{V_i}$$

This is called the open-loop gain.

This is the gain of the amplifier without feedback.

The feedback network extracts a voltage $V_f = \beta V_0$ from output V_0 of the amplifier. This voltage is added (positive feedback) or subtracted (negative feedback) from the signal voltage V_S. Now

$$V_i = V_S + V_f = V_S + \beta V_0 \text{ (Positive feedback)}$$

$$V_i = V_S - V_f = V_S - \beta V_0 \text{ (Negative feedback)}$$

The quantity $\beta = V_f/V_0$ is called feedback ratio or feedback fraction.

Gain of the amplifier without feedback, $A = V_0/V_i$.

or $$V_0 = AV_i \quad \text{...(1)}$$

Let us assume that

$$V_i = V_S + \beta V_0 \text{ (positive feedback)} \quad \text{...(2)}$$

From Eqs. (1) and (2), we get

$$V_0 = A[V_S + \beta V_0] = AV_S + A\beta V_0$$

or $$V_0[1 - A\beta] = AV_S$$

$$\therefore \quad \frac{V_0}{V_S} = \frac{A}{1 - A\beta} \quad \text{...(3)}$$

Let A_f be the overall gain (gain with feedback) of the amplifier. This is defined as the ratio of output voltage V_0 to the applied signal voltage V_S.

The resultant gain of the amplifier with feedback

$$A_f = \frac{\text{Output voltage}}{\text{Input signal voltage}} = \frac{V_0}{V_S} \quad \text{...(4)}$$

A_f is called the *closed-loop gain*.

From Eqs. (3) and (4)

$$A_f = \frac{A}{1 - A\beta} \quad \text{positive feedback} \quad \text{...(5)}$$

$$A_f = \frac{A}{1 + A\beta} \quad \text{negative feedback} \quad \text{...(6)}$$

(*i*) For positive feedback, $A_f > A$.

The positive feedback thus increases the gain of the amplifier.

(*ii*) For negative feedback, $A_f < A$.

The negative feedback thus decreases the gain of the amplifier.

(*iii*) In the case of positive feedback, if $\beta A = 1$, then $A_f \to \infty$.

This means that the amplifier becomes an oscillator.

Hence, the conditions for the amplifier to be an oscillator are

(*i*) *the feedback must be positive* and

(*ii*) *the feedback factor must be unity i.e.*, $\beta A = 1$.

74.3 Advantages of Negative Feedback

(*i*) Gain Stability : The gain of the amplifier employing negative feedback is given by

$$A_f = \frac{A}{1+\beta A} \qquad \text{...(1)}$$

If $\beta A >> 1$, then above equation approximates to,

$$A_f = \frac{A}{1+\beta A} \approx \frac{1}{\beta} \qquad \text{...(2)}$$

The term β is the feedback ratio and it is a function of the feedback network only. It does not depend upon the amplifier parameter changes or supply voltage variations. Thus the negative feedback improves stability of gain of the amplifier.

(*ii*) Reduction of distortion : The distortion in the output is reduced, because there is partial cancellation of out-of-phase signals. Both amplitude distortion and frequency distortion are reduced.

(*iii*) Reduction of output noise : The output noises and hum in an amplifier are reduced by negative feedback.

(*iv*) Increase in input impedance : Negative feedback increases the input impedance of an amplifier, which will not then draw much current from the source.

(*v*) Decrease in output impedance : Application of negative series voltage feedback in an amplifier *reduces the output impedance* which is a desirable characteristic for an amplifier because an amplifier with low output impedance can deliver power (or voltage) to the load without much loss.

(*vi*) Increase in band width : Negative feedback decreases the lower cut-off frequency but increases the upper cut-off frequency. Thus negative feedback increases the band width. But this increase in band width is achieved at the cost of reduction in the gain of the amplifier as shown in Fig. 74.2.

If A is the gain of amplifier without feedback and A_f with negative feedback, then since gain-band width product remains the same

$$A \times BW = A_f \times BW'$$

or

$$A \times (f_2 - f_1) = A_f \times (f'_2 - f'_1)$$

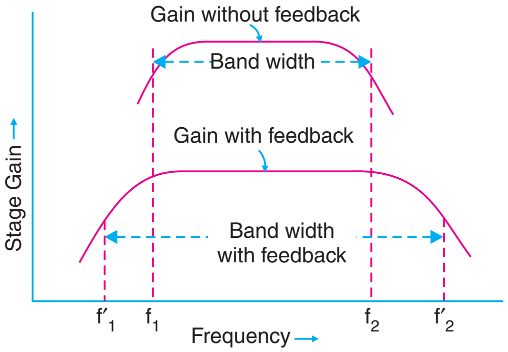


Fig. 74.2

74.4 Practical Feedback Circuits

1. Current Negative Feedback Transistor Amplifier

In these circuits, the voltage feedback to the input of the amplifier is proportional to the output current.

Circuit Details

Consider the circuit shown in Fig. 74.3. It is an $R - C$ coupled amplifier without emitter by-pass capacitor across emitter resistance R_E. R_1 and R_2 provide the base bias.

Circuit Operation

When a signal is applied at the input, there will be change in emitter current. The a.c. emitter current i_e produces a voltage across the emitter resistance R_E and this voltage $= i_e R_E$ with emitter becoming more positive during the increase of the signal voltage. Effectively, the value of v_{be} is decreased *i.e.*, the voltage developed across R_E is in series with and opposing the applied signal, v_s. This is negative feedback and it is proportional to the current through the emitter resistance. Thus the circuit works as a *negative current feedback* amplifier.

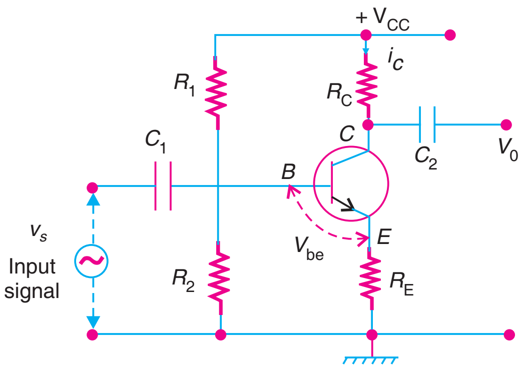


Fig. 74.3

Expression for closed loop gain

The voltage $i_e R_E$ across emitter resistance opposes the signal voltage v_s as it is in opposite direction to the signal.

The net input to the amplifier after negative feedback is

$$v_{be} = v_s - i_e R_E$$

$$v_{be} = v_s - i_c R_E \text{ since } i_e = i_c \qquad ...(1)$$

Output voltage v_0 is given by

$$v_0 = -i_c R_c$$

$$\therefore \quad i_c = -v_0/R_c \qquad ...(2)$$

Putting this value in Eq. (1), we get

$$v_{be} = v_s + v_0 \cdot \frac{R_E}{R_C}$$

$$v_{be} = v_s + v_0\beta \qquad \text{where } \beta = R_E/R_C \text{ is the feedback ratio.}$$

$$\frac{v_{be}}{v_0} = \frac{v_s}{v_0} + \beta$$

$$\frac{1}{A_f} = \frac{1}{A} + \beta$$

since the closed loop gain $A_f = v_0/v_{be}$ and the open loop gain $A = v_0/v_s$.

$$\therefore \quad A_f = \frac{A}{1 + A\beta} \qquad ...(3)$$

But since $\beta A >> 1$, the above expression approximates to

$$A_f \approx \frac{A}{\beta A} = \frac{1}{\beta} = \frac{R_C}{R_E} \qquad ...(4)$$

$\therefore$ The closed loop gain of the negative feedback amplifier is

$$A_f = \left(\frac{R_C}{R_E}\right)$$

2. Voltage–Series Feedback FET Amplifier

Fig. 74.4 shows an FET amplifier stage with voltage-series feedback. A part of the output signal (V_0) is obtained using a feedback network of resistors R_1 and R_2. The feedback voltage V_f is connected in series with the source signal V_S.

Their difference is the input signal V_i.

Without feedback the amplifier gain is

$$A = \frac{V_0}{V_i} = -g_m R_L \qquad ...(1)$$

Here R_L is the parallel combination of resistors:

$$R_L = R_D R_0 (R_1 + R_2) \qquad ...(2)$$

The feedback network provides a feedback factor of

$$\beta = \frac{V_f}{V_0} = \frac{-R_2}{R_1 + R_2} 2 \qquad ...(3)$$

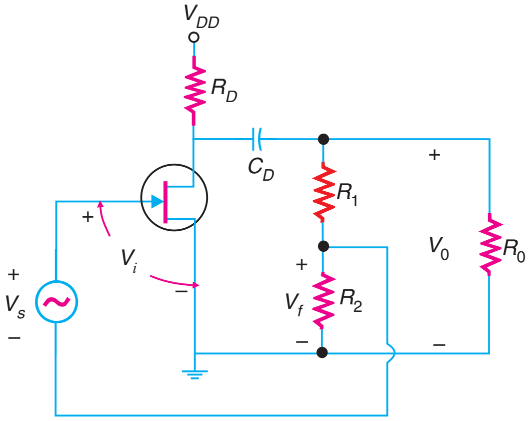


Fig. 74.4

The gain with negative feedback for the FET amplifier is

$$A_f = \frac{A}{1+\beta A} = \frac{-g_m R_L}{1+[R_2 R_L/(R_1+R_2)]g_m} \qquad ...(4)$$

If $\beta A >> 1$, we have

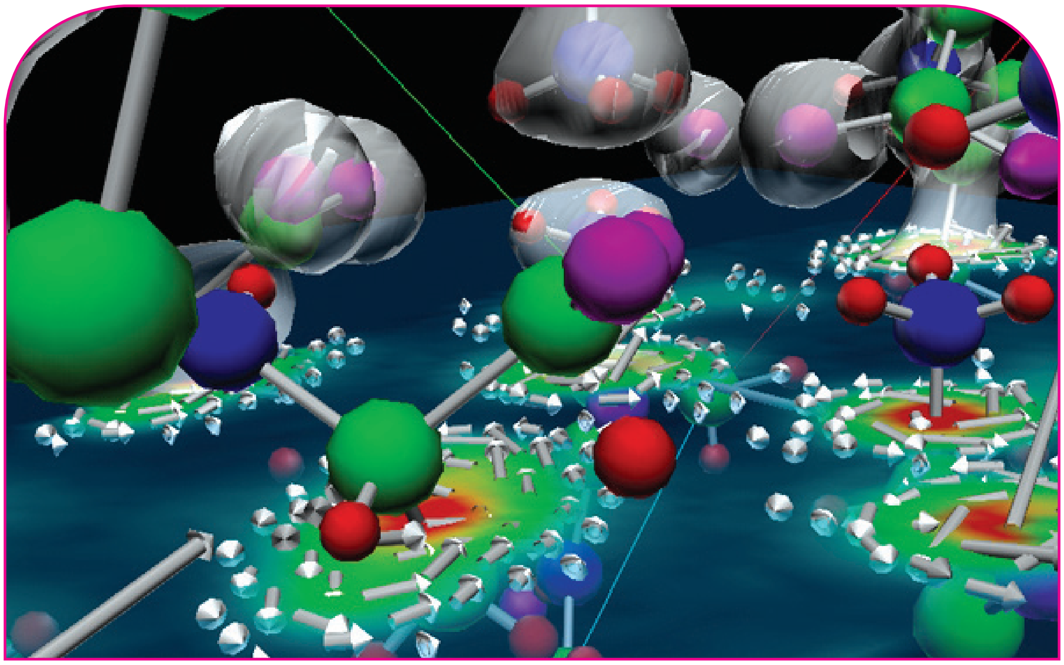
$$\boxed{A_f \cong \frac{1}{\beta} = -\frac{R_1 + R_2}{R_2}} \qquad ...(5)$$

PART – IX

STATISTICAL MECHANICS

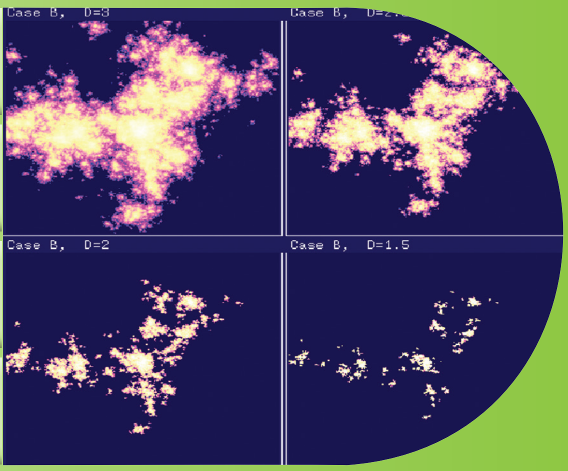
75. Statistical Mechanics I

76. Statistical Mechanics II



CHAPTER

75



STATISTICAL MECHANICS I

AT A GLANCE

75.1	Introduction
75.2	Phase Space
75.3	Maxwell-Boltzmann Distribution Law
75.4	Molecular Energies in an Ideal Gas
75.5	Bose-Einstein Distribution Law
75.6	Fermi-Dirac Distribution Law
75.7	Comparison of the Three Distribution Laws
75.8	Black-Body Radiation
75.9	Rayleigh-Jeans Formula
75.10	Planck Radiation Formula
75.11	Wien's Displacement Law
75.12	Stefan-Boltzmann Law from Planck's Formula

75.1 Introduction

Consider an isolated system composed of a large number N of particles, in which each particle has available to it several states with energies E_1, E_2, E_3... At a particular time the particles are distributed among the different states, so that n_1 particles have energy E_1 ; n_2 particles have energy E_2; and so on. The total number of particles is

$$N = n_1 + n_2 + n_3 + \ldots = \sum_i n_i \qquad \ldots(1)$$

We assume that the total number of particles remains constant for all processes occuring in the system.

The total energy of the system is

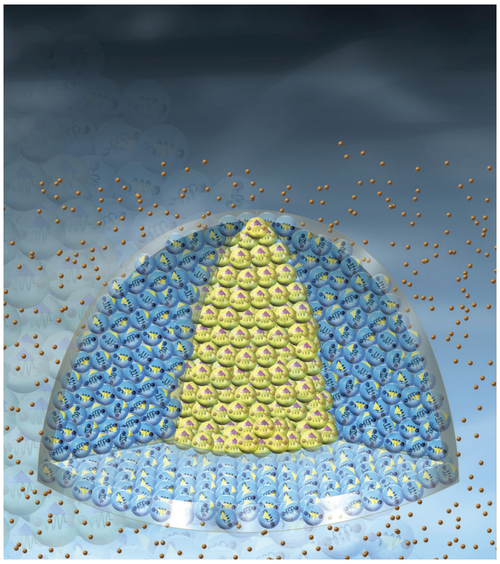
$$E = n_1 E_1 + n_2 E_2 + \ldots = \sum_i n_i E_i \qquad \ldots(2)$$

If the system is isolated, the total energy E must be constant. According to the basic postulate of statistical mechanics, the greater the number W of different ways in which the particles can be arranged among the available states to yield a particular distribution of energies, the more probable is the distribution. It is assumed that each state of a certain energy is equally likely to be occupied.

Therefore, the procedure for determining the most probable statistical distribution involves two steps: (*i*) obtaining the number of distinguishable arrangements (W) which give rise to the same distribution, and (*ii*) maximizing this number of arrangements (W) with respect to different distributions.

We shall consider systems of three different kinds of particles.

1. Maxwell-Boltzmann Statistics: This is applicable to the identical, distinguishable particles of any spin. The molecules of a gas are the particles of this kind.



Metallic Phases for Bosons.

2. Bose-Einstein Statistics: This is applicable to the identical, indistinguishable particles of zero or integral spin. These particles are called *bosons.* The examples of bosons are helium atoms at low temperature and the photons.

3. Fermi-Dirac Statistics: This is applicable to the identical indistinguishable particles of half-integral spin. These particles obey Pauli exclusion principle and are called *fermions.* The examples of fermions are electrons, protons, neutrons, etc.

75.2 Phase Space

To specify the position as well as energy of a molecule inside a gas, we must specify three space coordinates x, y, z and three momentum coordinates p_x, p_y, p_z. As a purely mathematical concept, we may imagine six dimensional space in which coordinates are x, y, z, p_x, p_y, p_z. This six dimensional space for a single molecule is called *phase space* or μ-*space*. The instantaneous state of a particle in the phase space is represented by a point known as phase point or representative point. In phase space we may consider an element of volume $dx, dy, dz, dp_x, dp_y, dp_z$. Any such element of volume in six dimensional space is called a *cell*. Thus the phase space may be divided into a large number of cells. A cell may contain a large number of phase points. Each cell in phase space has the volume h^3.

75.3 Maxwell-Boltzmann Distribution Law

Consider a system of N distingishable molecules of a gas. Suppose N_1 of them have energy E_1, n_2 have energy E_2, ... n_i have energy E_i and so on. Thus the entire assembly of molecules can be divided into different energy states with energies $E_1, E_2, E_3, \ldots E_i$ and having $n_1, n_2, n_3, \ldots n_i$ molecules.

(1) The total number of molecules N is constant. Hence

$$N = n_1 + n_2 + n_3 + \ldots + n_i + \ldots = \text{constant}$$

$$\text{or} \quad \delta N = \delta n_1 + \delta n_2 + \delta n_3 + \ldots + \delta n_i + \ldots = 0$$

i.e., $$\sum_i \delta n_i = 0 \qquad \text{...(1)}$$

(2) The total energy E of the gas molecules is constant. Hence

$$E = E_1 n_1 + E_2 n_2 + E_3 n_3 + \ldots + E_i n_i + \ldots = \text{constant}$$

or $\delta E = E_1 \delta n_1 + E_2 \delta n_2 + E_3 \delta n_3 + \ldots + E_i\, \delta n_i + \ldots = 0$

i.e., $$\sum_i E_i\, \delta n_i = 0 \qquad \text{...(2)}$$

(3) Suppose there are g_i cells with the energy E_i. The total number of ways in which n_i molecules can have the energy E_i is $(g_i)^{n_i}$. Hence the total number of ways in which N molecules can be distributed among the various energies is

$$W_1 = (g_1)^{n_1} (g_2)^{n_2} (g_3)^{n_3} \ldots (g_i)^{n_i} \ldots$$

The number of ways in which the groups of n_1, n_2, n_3 ... n_i ... particles can be chosen from N particles is given by

$$W_2 = \frac{N!}{n_1!\, n_2!\, n_3! \ldots}$$

The number of distinguishable ways in which N molecules can be distributed among the possible energy levels is

$$W = W_1 W_2 = \frac{N!}{n_1!\, n_2!\, n_3! \ldots} (g_1)^{n_1} (g_2)^{n_2} (g_3)^{n_3} \ldots \qquad \text{...(3)}$$

The quantity W is called the *thermodynamic probability* for the system. For the most probable distribution, W is a maximum subject to the restriction that the total number of particles N and the total energy E are constants.

The natural logarithm of Eq. (3) is

$\ln W = \ln N! - \Sigma \ln n_i! + \Sigma n_i \ln g_i$

By Stirling's theorem, $1\text{n}\, n! = n \ln n - n$

$\ln W = N \ln N - N - \Sigma\, n_i \ln n_i + \Sigma\, n_i + \Sigma\, n_i \ln g_i$

$\ln W = N \ln N - \Sigma\, n_i \ln n_i + \Sigma\, n_i \ln g_i$ ...(4)

From Eq. (4), we have for maximum W

$$\delta \ln W_{\max} = -\Sigma n_i \frac{1}{n_i} \delta n_i - \Sigma (\ln n_i)\, \delta n_i + \Sigma (\ln g_i)\, \delta n_i = 0$$

or $-\Sigma (\ln n_i)\, \delta n_\text{i} + \Sigma (\ln g_\text{i})\, \delta n_i = 0$ $[\because \Sigma\, \delta n_i = 0]$ ...(5)

Eqs. (1) and (2) can be incorporated into Eq. (5) by making use of Lagrange's method of undetermined multipliers. Multiplying Eq. (1) by $-\alpha$ and Eq. (2) by $-\beta$ and adding to Eq. (5), we get

$\Sigma (-\ln n_i + \ln g_i - \alpha - \beta E_i)\, \delta n_i = 0$ ...(6)

or $-\ln n_i + \ln g_i - \alpha - \beta E_i = 0$

or $$n_i = g_i\, e^{-\alpha}\, e^{-\beta E_i} \qquad \text{...(7)}$$

Eq. (7) is called Maxwell-Boltzmann *distribution law*.

M-B distribution in terms of temperature. It can be shown that $\beta = 1/kT$ where k = Boltzmann's constant and T = absolute temperature of the gas.

$\therefore$ $$n_i = g_i\, e^{-\alpha}\, e^{-(E_i/kT)}. \qquad \text{...(8)}$$

75.4 Molecular Energies in an Ideal Gas

The M.B. distribution law is

$$n_i = g_i e^{-\alpha} e^{-(E_i/kT)} \qquad \text{...(1)}$$

Consider an ideal gas that contains N molecules. Consider a continuous distribution of molecular energies instead of the discrete set E_1, E_2, E_3, ... Then Eq. (1) becomes

$$n(E)\, dE = g(E)\, e^{-\alpha} e^{-(E/kT)}\, dE \qquad \text{...(2)}$$

Here $n(E)\, dE$ represents the number of molecules having energies between E and $E + dE$. g (E) dE represents the number of states that have energies between E and $E + dE$.

Let us first find $g(E)\, dE$. A molecule of energy E has a momentum p whose magnitude is

$$p = \sqrt{2mE} = \sqrt{P_x^2 + P_y^2 + P_z^2} \qquad \text{...(3)}$$

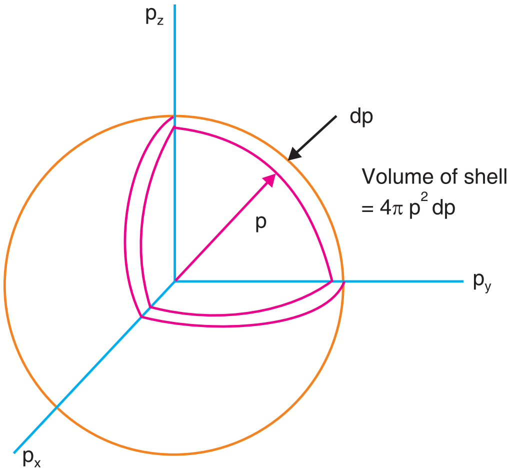


Fig. 75.1

Each set of momentum components p_x, p_y, p_z specifies a different state of motion. Let us imagine a *momentum space* whose coordinate axes are p_x, p_y, p_z (Fig. 75.1). The number of momentum states available to a particle with a momentum whose magnitude is between p and $p + dp$ is proportional to the volume of a spherical shell in momentum space of radius p and thickness dp. Volume of this spherical shell is $4\pi p^2\, dp$. Hence

$$g(p)\, dp = B\, p^2\, dp \qquad \text{...(4)}$$

where B is some constant. Since each momentum magnitude p corresponds to a single energy E, the number of energy states $g(E)\, dE$ between E and $E + dE$ is the same as the number of momentum states $g(p)\, dp$ between p and $p + dp$.

$$\therefore \qquad g(E)\, dE = Bp^2\, dp \qquad \text{...(5)}$$

But $p^2 = 2mE$ and $dp = \dfrac{m\, dE}{\sqrt{2mE}}$

$$\therefore \qquad g(E)\, dE = \sqrt{2}\, m^{3/2} B \sqrt{E}\, dE \qquad \text{...(6)}$$

Substituting for $g(E)\, dE$ in Eq. (2),

$$n(E)\, dE = C\sqrt{E}\, e^{-(E/kT)}\, dE \qquad \text{...(7)}$$

where $\qquad C = \sqrt{2}\, m^{3/2}\, Be^{-\alpha} = a$ constant.

To find C we make use of the normalization condition that the total number of molecules is N. Therefore,

$$N = \int_0^\infty n(E)\, dE = C\int_0^\infty \sqrt{E}\, e^{-(E/kT)}\, dE \qquad \text{...(8)}$$

From a table of definite integrals we find that

$$\int_0^\infty \sqrt{x}\, e^{-ax}\, dx = \frac{1}{2a}\sqrt{\frac{\pi}{a}}$$

$$\therefore \qquad N = C\frac{1}{2(1/kT)}\sqrt{\frac{\pi}{(1/kT)}} = C \times \frac{\sqrt{\pi}}{2} \times (kT)^{3/2}$$

or
$$C = \frac{2\pi N}{(\pi kT)^{3/2}} \qquad ...(9)$$

∴
$$n(E)\,dE = \frac{2\pi N}{(\pi kT)^{3/2}} \sqrt{E}\, e^{-E/kT}\, dE \qquad ...(10)$$

Eq. (10) is plotted in Fig. 75.2 in terms of kT.

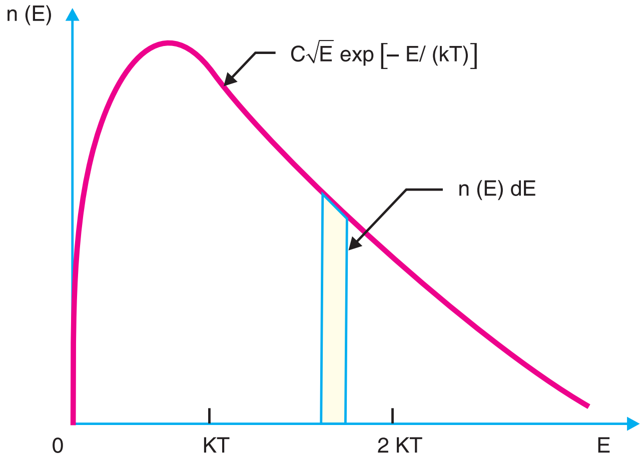


Fig. 75.2

The total internal energy of the system is

$$E = \int_0^\infty E\,n(E)\,dE \qquad \frac{}{(\ kT)^{/}} \int_0^\infty E^{/}\ e^{\ E/kT}\, dE$$

The value of the definite integral is $\frac{3}{4}(kT)^2\sqrt{\pi kT}$.

∴
$$E = \frac{2\pi N}{(\pi kT)^{3/2}} \times \frac{3}{4}(kT)^2\sqrt{\pi kT} = \frac{3}{2}NkT \qquad ...(11)$$

The average energy of an ideal-gas molecule is E/N.

∴
$$\overline{E} = \frac{3}{2}kT\ . \qquad ...(12)$$

Maxwell-Boltzmann Velocity Distribution Law

Substitute $E = 1/2\ mv^2$ and $dE = mv\ dv$ in Eq. (10). Then we get

$$n(v)\,dv = \frac{\sqrt{2}\,\pi N m^{3/2}}{(\pi kT)^{3/2}} v^2 e^{-mv^2/2kT}\, dv \qquad ...(13)$$

This equation represents the number of molecules with speeds between v and $v + dv$ in an assembly of ideal gas containing N molecules at absolute temperature T. This formula is plotted in Fig. 75.3.

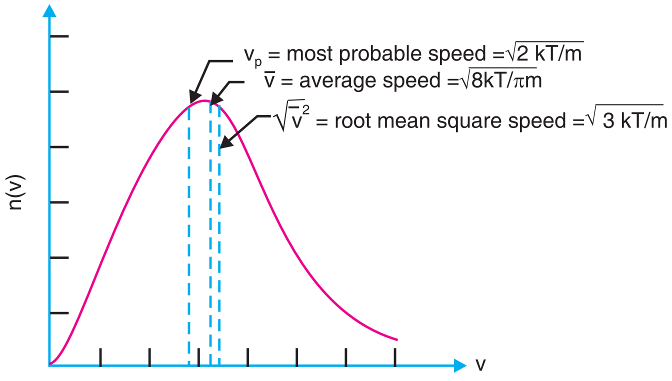


Fig. 75.3

75.5 Bose-Einstein Distribution Law

Consider an assembly of N bosons. They are identical and indistinguishable. No restriction is imposed as to the number of particles that may occupy a given cell. Let us now consider a box divided into g_i sections by $(g_i - 1)$ partitions and n_i indistinguishable particles to be distributed among these sections. The permutations of n_i particles and $(g_i - 1)$ partitions simultaneously is given by $(n_i + g_i - 1)!$. But this includes also the permutations of n_i particles among themselves and also $(g_i - 1)$ partitions among themselves, as both these groups are internally indistinguishable. Hence the actual number of ways in which n_i particles are to be distributed in g_i sublevels is

$$\frac{(n_i + g_i - 1)!}{n_i!(g_i - 1)!}.$$

Therefore, the total number of distinguishable and distinct ways of arranging N particles in all the variable energy states is given by

$$W = \Pi \frac{(n_i + g_i - 1)!}{n_i!(g_i - 1)!} \qquad ...(1)$$

n_i and g_i are large numbers. Hence we may neglect 1 in the above expression.

$$\therefore \qquad W = \Pi \frac{(n_i + g_i)!}{n_i!\,g_i!} \qquad ...(2)$$

$$\ln W = \Sigma\,[\ln (n_i + g_i)\,! - \ln n_i\,! - \ln g_i\,!]$$

As n_i and g_i are large numbers, we can use Stirling approximation.

$$\ln W = \Sigma\,(n_i + g_i) \ln (n_i + g_i) - n_i \ln n_i - g_i \ln g_i \qquad ...(3)$$

Here, g_i is not subject to variation and n_i varies continuously.

For most probable distribution, $\delta \ln W_{\text{max}} = 0$.

Hence, if the W of Eq. (3) represents a maximum,

$$\delta \ln W_{max} = \Sigma\,[\ln (n_i + g_i) - \ln n_i]\,\delta n_i = 0 \qquad ...(4)$$

The total number of particles and total energy are constants.

$$\therefore \qquad \Sigma\,\delta n_i = 0 \qquad ...(5)$$

$$\Sigma\,E_i\,\delta n_i = 0 \qquad ...(6)$$

Multiplying Eq. (5) by $-\alpha$, Eq. (6) by $-\beta$ and adding to Eq. (4),

$$\Sigma\,[\ln (n_i + g_i) - \ln n_i - \alpha - \beta E_i]\,\delta n_i = 0$$

The variations δn_i are independent of each other. Hence we get

$$\ln\left(\frac{n_i + g_i}{n_i}\right) - \alpha - \beta E_i = 0$$

or

$$n_i = \frac{g_i}{(e^{\alpha + \beta E_i}) - 1} \qquad ...(7)$$

or

$$n_i = \frac{g_i}{(e^{\alpha}\,e^{E_i/kT}) - 1} \qquad ... (8)$$

This is known as Bose-Einstein distribution law.

75.6 Fermi-Dirac Distribution Law

F-D statistics is obeyed by indistinguishable particles of half-integral spin that have antisymmetric wave functions and obey Pauli exclusion principle. Consider N ferminons with the total energy E. Suppose that n_i particles occupy the first energy level with energy E_1, n_2 particles occupy the second energy level with energy E_2 and so on. Let us now find out the total number of ways in which n_i particles can be distributed in g_i cells having the same energy E_i.

The number of distinguishable arrangements of n_i particles in g_i cells is

$$\frac{g_i\,!}{n_i\,!(g_i - n_i)!}$$

The total number of Eigen states for the whole system is given by

$$W = \Pi \frac{g_i\,!}{n_i\,!(g_i - n_i)!} \qquad \text{... (1)}$$

Taking the natural logarithm of both sides,

$$\ln W = \Sigma\,[\ln g_i\,! - \ln n_i\,! - \ln (g_i - n_i)\,!]$$

Applying Stirling's approximation,

$$\ln W = \Sigma\,[g_i \ln g_i - g_i - n_i \ln n_i + n_i - (g_i - n_i) \ln (g_i - n_i) + (g_i - n_i)]$$

$$\ln W = \Sigma\,[g_i \ln g_i - n_i \ln n_i - (g_i - n_i) \ln (g_i - n_i)] \qquad \text{...(2)}$$

Here g_i is not subject to variation and n_i varies continuously.

For most probable distribution, $\delta \ln W_{max} = 0$.

$$\delta \ln W_{max} = \Sigma\,[-\ln n_i + \ln (g_i - n_i)]\,\delta n_i = 0 \qquad \text{... (3)}$$

But $\qquad \Sigma\,\delta n_i = 0 \qquad$ ... (4)

and $\qquad \Sigma\,E_i\,\delta n_i = 0 \qquad$ ... (5)

Multiplying Eq. (4) by $-\alpha$, Eq. (5) by $-\beta$ and adding to Eq. (3),

$\Sigma\,[-\ln n_i + \ln (g_i - n_i) - \alpha - \beta\,E_i]\,\delta\,n_i = 0$

As the variations $\delta\,n_i$ are independent of each other, we get

$$\ln\left(\frac{g_i - n_i}{n_i}\right) - \alpha - \beta E_i = 0$$

or

$$n_i = \frac{g_i}{(e^{\alpha}\,e^{\beta E_i}) + 1} \qquad \text{...(6)}$$

This is known as Fermi-Dirac Distribution law.

$\beta = 1/kT$ and $\alpha = -E_F/kT$ where E_F is called Fermi energy. $E_F = -\,\alpha kT$.

$$n_i = \frac{g_i}{(e^{\alpha}\,e^{E_i/kT}) + 1}\,. \qquad \text{...(7)}$$

Definition of Fermi energy : *At the absolute zero of temperature the maximum kinetic energy that a free electron can have is called the Fermi energy, E_F.*

$$\therefore \qquad n_i = \frac{g_i}{\left[e^{(E_i - E_F)/kT}\right] + 1}\,.$$

75.7 Comparison of the Three Distribution Laws

Let us define a quantity $f(E_i) = \dfrac{n_i}{g_i}$.

$f(E_i)$ is called the 'occupation index' of a state of energy E_i.

$$f_{MB}\,(E_i) = \frac{1}{e^{\alpha}\,e^{E_i/kT}}$$

$$f_{BE}(E_i) = \frac{1}{(e^{\alpha} e^{E_i/kT}) - 1}$$

$$f_{FD}(E_i) = \frac{1}{(e^{\alpha} e^{E_i/kT}) + 1}$$

When $E >> kT$, f_{BE} and f_{FD} approach f_{MB}. In general,

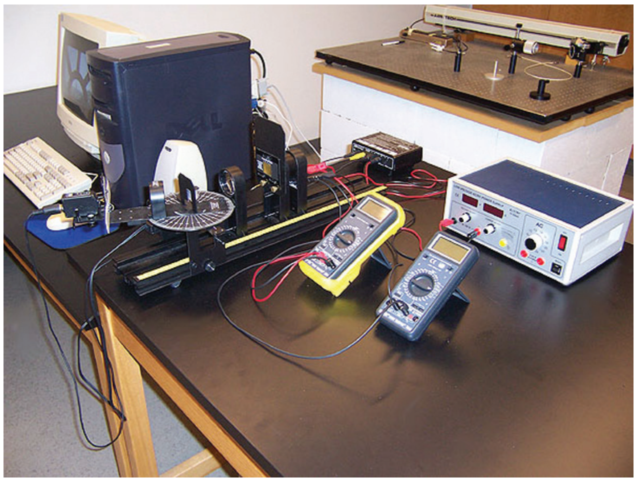
$$f(E) = \frac{1}{\exp[\alpha + E/(kT)] + \delta}$$

is the probability of occupation of a single state at energy E and δ is +1 for F-D, 0 for M-B, and – 1 for B-E statistics. We will refer to $f(E)$ as the *occupation probability* or the *distribution function.*

Comparison of MB, BE, and FD Distribution Functions

MB	*BE*	*FD*
(1) Holds for distinguishable particles; approximations of BE and FD distributions at $E >> kT$. (2) $f(E) = Ae^{-E/kT}$ (3) Applies to common gases at normal temperatures.	Holds for indistinguishable particles not obeying Pauli's exclusion principle $f(E) = \dfrac{1}{e^{\alpha} e^{E/kT} - 1}$ Applies to photon gas, phonon gas.	Holds for indistinguishable particles obeying the exclusion principle. $f(E) = \dfrac{1}{e^{(E-E_F)/kT} + 1}$ Applies to electron gas in metals.

75.8 Black-Body Radiation



Black-Body Radiation.

The first indication of the inadequacy of classical ideas to explain the properties of matter, occurred in *black-body radiation*. A black-body is a body which absorbs all the radiation incident on it, and hence is the perfect absorber. Consideration of equilibrium of different bodies at the same temperature implies that it is also the best emitter of radiation energy. A blackbody may be idealized by a small hole drilled in a cavity. If the radiation from a black-body is analysed by a spectrometer, it is found that the intensity distribution as a function of frequency has a well-defined shape. The spectrum of blackbody radiation is shown in Fig. 75.4 for two temperatures. The spectral distribution of energy in the radiation depends only on the temperature of the body. What is most significant is that, for a given temperature, it is a universal curve independent of the properties of the walls of the cavity.

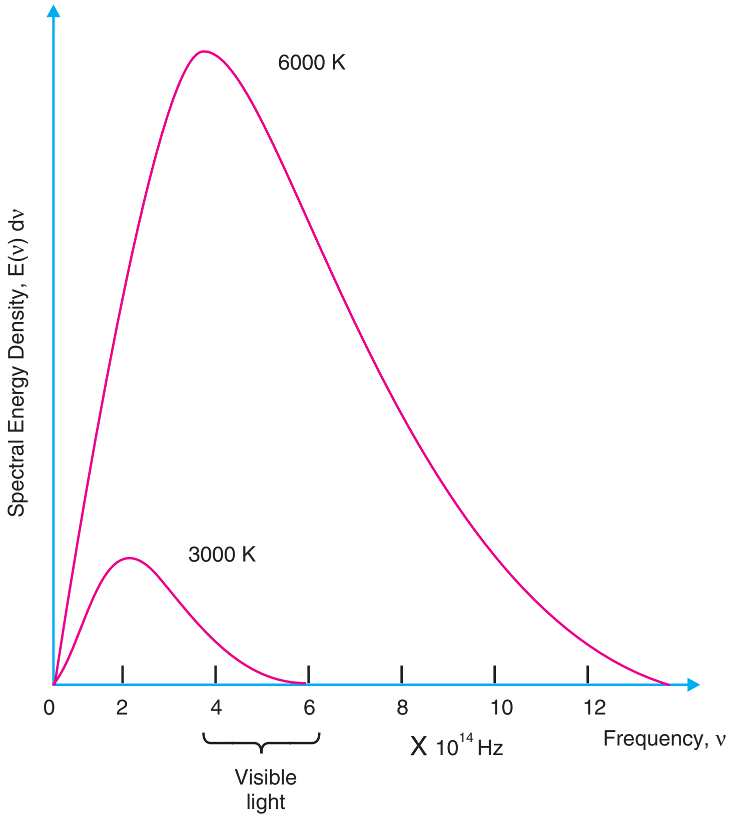


Fig. 75.4

75.9 Rayleigh-Jeans Formula

Rayleigh's formula for the distribution of energy in the black-body spectrum is based on the principle of equipartition of energy for all the possible modes of free vibration which might be assigned to radiation. Thus they considered average energy of an oscillator (*i.e.*, per mode of vibration) as

$$\bar{\varepsilon} = kT \text{ (classical result).}$$

The number of independent standing waves per unit volume in a cavity in the frequency range ν and $\nu + d\nu$ is given by $G(\nu)\,d\nu \ \dfrac{8\pi\nu^2}{c^3}d\nu$.

The energy $u(\nu)\,d\nu$ per unit volume in the cavity within frequency range ν and $\nu + d\nu$ is,

$$u(\nu)d\nu = \bar{\varepsilon}\,G(\nu)\,d\nu = kT \times \frac{8\pi\nu^2}{c^3}d\nu = \frac{8\pi\nu^2\,kT\,d\nu}{c^3}$$

Rayleigh Jean's law agrees well with the experimental results at low frequencies but near the maximum in the spectrum and at higher frequencies it is in violent disagreement (Fig. 75.5). According to this law, the energy density will continuously increase with increase in frequency ν and approaches ∞ when $\nu \to \infty$.

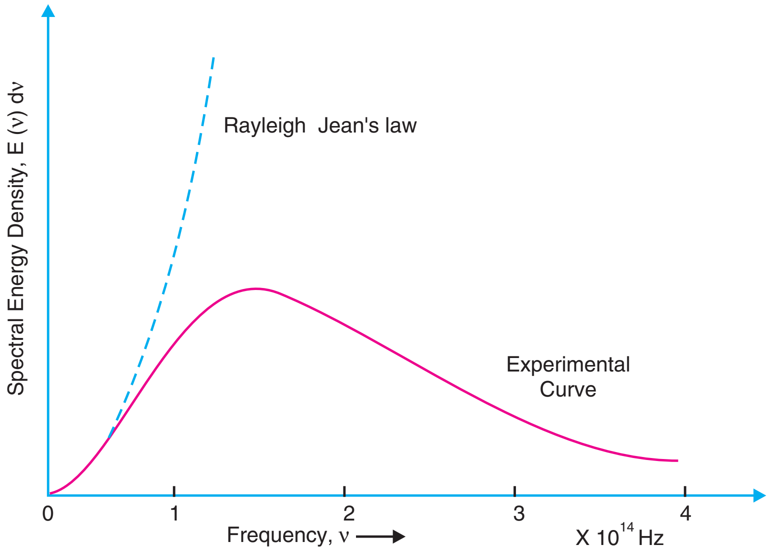


Fig. 75.5

This is in contradiction to observed results. This fatal objection to the law has been known as ultraviolet catastrophe.

Furthermore, at any temperature T the total energy of radiation, $\int_0^\infty \frac{8\pi\nu^2 kT\,d\nu}{c^3}$, is predicted to be infinite which is against the Stefan's law.

75.10 Planck Radiation Formula

We shall study the application of the B-E distribution to photons and obtain Planck's radiation formula. Consider a box containing electromagnetic radiation (*i.e.*, photons) at a temperature T. These photons would be absorbed and re-emitted by the container walls. Under thermal equilibrium, the energy distribution can be determined by the Bose-Einstein statistics. However, there is an important difference between the photons and the other particles, namely, the photon number is not conserved. This is because the walls can absorb a photon of energy E and emit two photons of energy E_1 and E_2 such that $E_1 + E_2 = E$. Thus, we have only the energy restriction and no restriction on the photon number.

The B. E. distribution function is

$$f(E) = \frac{1}{(e^{\alpha}\, e^{E/kT}) - 1} \qquad \text{...(1)}$$

The non-conservation of photons means that $\alpha = 0$.

Therefore, the B-E distribution function for photons is

$$f(\nu) = \frac{1}{e^{h\nu/kT} - 1} \quad (\because E = h\nu) \qquad \text{...(2)}$$

The number of independent standing waves per unit volume in the frequency interval from ν to $\nu + d\nu$ is

$$G(\nu)\,d\nu = \frac{8\pi\nu^2}{c^3}\,d\nu \qquad \text{...(3)}$$

The energy-density of radiation in the frequency range ν to $\nu + d\nu$ is

$$u(\nu)\,d\nu = h\nu\,G(\nu)\,f(\nu)\,d\nu$$

$$= \frac{8\pi h}{c^3}\left(\frac{\nu^3\,d\nu}{e^{h\nu/kT}-1}\right) \qquad \text{...(4)}$$

This is Planck radiation formula. Planck's formula agrees with experimental curves.

75.11 Wien's Displacement Law

Eq. (4) can be expressed in terms of wavelength of radiation as follows:

$$u_\lambda\,d\lambda = \frac{8\pi ch}{\lambda^5}\left(\frac{1}{e^{hc/\lambda kT}-1}\right)d\lambda$$

or
$$u_\lambda = 8\pi hc(\lambda^{-5})\left[\exp.\left(\frac{hc}{\lambda kT}\right)-1\right]^{-1}.$$

Let $\lambda_{\max}$ be the wavelength whose energy density is the greatest. Then,

$$\frac{du_\lambda}{d\lambda} = 0$$

Hence we get,
$$\lambda_{\max} T = \frac{hc}{4.965k} = 2.898\times10^{-3}\text{ mK}.$$

This is Wien's displacement law.

The peak in the blackbody spectrum shifts to progressively shorter wavelengths (higher frequencies) as the temperature is increased (Fig. 75.4).

75.12 Stefan-Boltzmann Law from Planck's Formula

Planck's radiation formula is

$$u(\nu)\,d\nu = \frac{8\pi h}{c^3}\frac{\nu^3 d\nu}{e^{h\nu/kT}-1}$$

The total energy density u over all frequencies is

$$u = \int_0^\infty u(\nu)\,d\nu = \int_0^\infty \frac{8\pi h\nu^3}{c^3}\frac{d\nu}{e^{h\nu/kT}-1} = aT^4$$

where a is a constant.

The energy R radiated by an object per second per unit area is

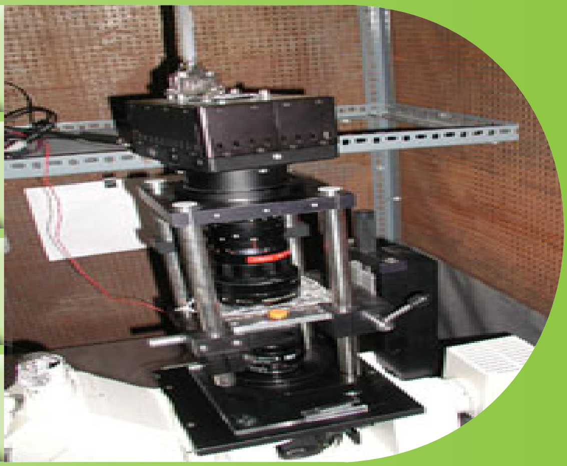
$$R = e\,\sigma\,T^4$$

σ = Stefan's constant = 5.670×10^{-8} $\text{Wm}^{-2}\,\text{K}^{-4}$.

The emissivity e depends on the nature of the radiating surface.

This is the Stefan-Boltzmann law.

CHAPTER 76



STATISTICAL MECHANICS II

AT A GLANCE

76.1	Macroscopic and Microscopic Descriptions
76.2	Ensembles
76.3	Probability
76.4	Thermodynamic Probability
76.5	Boltzmann's Theorem on Entropy and Probability
76.6	Fundamental Postulates of Statistical Mechanics
76.7	Statistical Equilibrium
76.8	Quantum Statistics
76.9	Electron Gas

76.1 Macroscopic and Microscopic Descriptions

We may mentally isolate any finite portion of matter from its local surroundings. We call such a portion the *system*. Everything outside the system which has a direct bearing on its behaviour we call the *environment*. The system and its environment taken together are usually referred to as the *universe*. Let us consider an example of a system and its surroundings. Consider some dilute gas of fixed mass enclosed in a cylinder provided with a frictionless piston. The gas is the system. The walls of the cylinder and the piston form the system boundary. The open atmosphere plays the role of the surroundings.

We want to determine the behaviour of the system by finding how it interacts with its environment. A system such as a dilute gas can be studied from (*i*) microscopic point of view or (*ii*) macroscopic point of view. Here the term '*micro*' stands for small scale and "*macro*" stands for large scale. Molecules form the '*micro*' systems, while any actual mass of gas is a '*macro*' system. A macroscopic system is thus made up of an assembly of a large number of microscopic ones. In the microscopic description, we consider quantities that describe the atoms and molecules that make up the system, their speeds, energies, masses, behaviour during collision etc. These quantities form the basis for the science of *statistical mechanics*. The variables used in microscopic description are (*i*) very large in number (*ii*) not perceptible to our organic senses and (*iii*) not easily measurable by simple experimental procedures. On the other hand, in the macroscopic description of a dilute gas, we focus our attention on the gas as a whole. The variables used in the macroscopic description are: (*i*) few in number (*ii*) perceptible to our senses and (*iii*) easily measurable by simple experimental procedures. *Macroscopic quantities* are gross properties of the system measured by laboratory operations. The collective behaviour of a macroscopic system may be represented by some parameters like pressure (P), volume (V), temperature (T), internal energy (U) and entropy (S). These parameters are always measurable through direct experiments. These macroscopic quantities form the basis for the science of *thermodynamics*. The microscopic interactions will interpret the properties of macroscopic system.

For any system, the macroscopic and the microscopic quantities must be related because they are simply different ways of describing the same situation. In particular, we should be able to express the macroscopic quantities in terms of the microscopic quantities. For example, a very large number of collisions of molecules with the walls give an average pressure P. Similarly, the temperature T of a gas may be related to the average kinetic energy of translation of the molecules. Thus, a very large number of position and momentum coordinates of all the molecules in microscopic description give only two macroscopic variables P and T. The macroscopic description is therefore, much simpler which is used in the analysis of thermodynamic properties of the systems. We can express the laws of thermodynamics quantitatively in the language of statistical mechanics.

76.2 Ensembles

A *system* is defined as a collection of a number of particles. *An ensemble is defined as a collection of a large number of macroscopically identical, but essentially independent systems.* By the term *macroscopically identical* we mean that each of the systems constituting an ensemble satisfies the same macroscopic conditions, *e.g.*, volume, energy, pressure, total number of particles etc. By the term *independent systems* we mean that the systems constituting an ensemble are mutually non-interacting. In an ensemble the systems play the same role as the non-interacting molecules do in a gas.

Brass Ensembles.

76.3 Probability

The probability of an event may be defined as the ratio of the number of cases in which the event occurs to the total number of cases. That is,

$$\text{the probability of the event} = \frac{\text{Number of cases in which the event occurs}}{\text{Total number of cases}}.$$

Suppose we toss a coin a large number of times and count the number of times the 'head' or 'tail' is uppermost. Experience shows that the 'heads' and 'tails' will be uppermost equal number of times. That is,

$$P\,(\text{head}) = 1/2 \text{ and } P\,(\text{tail}) = 1/2.$$

76.4 Thermodynamic Probability

The thermodynamic probability of a particular macrostate is defined as the number of microstates corresponding to that macrostate.

It is represented by W.

Consider two cells in phase space represented by i and j and four molecules a, b, c and d. Let N_i and N_j be the number of molecules in the cells i and j respectively. Then the possible macrostates are five in number (Fig. 76.1).

N_i	4	3	2	1	0
N_j	0	1	2	3	4

Fig. 76.1

In general, to each of these five macrostates there corresponds a different number of microstates. Let us consider the microstates corresponding to the macrostate $N_i = 3$, $N_j = 1$. The number of microstates corresponding to the macrostate $N_i = 3$, $N_j = 1$ is four, as shown in Fig. 76.2.

Cell i	abc	abd	acd	bcd
Cell j	d	c	b	a

Fig. 76.2

Therefore, the thermodynamic probability for the macrostate, $N_i = 3$, $N_j = 1$ is 4. That is $W = 4$.

76.5 Boltzmann's Theorem on Entropy and Probability

Boltzmann discovered a relation between entropy (a thermodynamical quantity) and probability (a statistical quantity). Boltzmann started from a simple idea that the equilibrium state of the system is the state of maximum probability. That is, the probability of the system in equilibrium state is maximum. But from the thermodynamic point of view, the equilibrium state of the system is the state of maximum entropy. If the system is not in equilibrium, then changes take place within the system until the equilibrium state or the state of maximum entropy is reached. Thus, in equilibrium state both the entropy and thermodynamical probability have their maximum values. As equilibrium state is the state of maximum entropy and maximum probability, Boltzmann concluded that entropy is a function of probability. That is,

$$S = f(W) \qquad \text{...(1)}$$

Here, S is entropy and W is the thermodynamical probability of the state. Let us consider two separate systems having entropies S_1 and S_2 and thermodynamic probabilities W_1 and W_2.

Then
$$S_1 = f(W_1) \text{ and } S_2 = f(W_2) \qquad \text{...(2)}$$

The total entropy of the two systems is

$$S_1 + S_2 = f(W_1) + f(W_2) \qquad \text{...(3)}$$

But the thermodynamic probability of the two systems taken together is $W_1 W_2$.

$$f(W_1 W_2) = f(W_1) + f(W_2) = S_1 + S_2 \qquad \text{...(4)}$$

If this relation is to be satisfied, $f(W)$ must be a logarithmic function of W.

$$\therefore \quad f(W) = k \log W$$

or
$$S = k \log W. \qquad \text{...(5)}$$

76.6 Fundamental Postulates of Statistical Mechanics

1. Any gas may be considered to be composed of molecules that are in motion and behave like very small elastic spheres.
2. All the cells in the phase space are of equal size.
3. All accessible microstates corresponding to possible macrostates are equally probable. This is called the *postulate of equal a priori probability.*
4. The equilibrium state of a gas corresponds to the macrostate of maximum probability.
5. *The total number of molecules is constant.*

76.7 Statistical Equilibrium

Boltzmann canonical principle is applied to determine the equilibrium state of the system. According to this principle, *Equilibrium state of a system is that which is most probable.*

Consider an isolated system composed of a large number N of particles, in which each particle has available to it several states with energies E_1, E_2, E_3, ... At a particular time the particles are distributed among the different states, so that n_1 particles have energy E_1; n_2 particles have energy E_2; and so on. The total number of particles is

$$N = n_1 + n_2 + n_3 + \ldots = \sum_i n_i \qquad \ldots(1)$$

We assume that the total number of particles remains constant for all processes occurring in the system.

The total energy of the system is

$$E = n_1 E_1 + n_2 E_2 + \ldots = \sum_i n_i E_i \qquad \ldots(2)$$

If the system is isolated, the total energy E must be constant. However, as a result of their mutual interactions and collisions, the distribution of the particles among the available energy states may be changing. For example in a gas, a fast molecule may collide with a slow one; after the collision, the fast molecule may have slowed down and the slow one may have sped up. Or an excited atom may collide inelastically with another atom, with a transfer of its excitation energy into kinetic energy of both atoms. Hence, in both examples, the particles after the collision are in different states. In other words, the numbers n_1, n_2, n_3, ... which give the distribution of the N particles among the available energy states, may be changing. It is reasonable to assume that for each macroscopic state of a system of particles there is a distribution which is more favoured than any other. In other words, we may say that *given the physical conditions of the system of particles*, *there is a most probable distribution.* When this distribution is achieved, the system is said to be in *statistical equilibrium*.

76.8 Quantum Statistics

Statistical mechanics can be divided into two main classes.

1. Classical Statistics or Maxwell-Boltzmann Statistics.
2. Quantum Statistics.

Classical statistics interpreted successfully many ordinarily observed phenomena such as temperature, pressure, energy etc. But it failed to account for several other experimentally observed phenomena such as black body radiation, photoelectric effect, specific heat capacity at low temperatures, etc. This failure of classical statistics forced the issue in favour of the new quantum idea of discrete exchange of energy between systems. Thus, a new quantum statistics was investigated. There are two types of quantum statistics :

1. Bose-Einstein statistics.
2. Fermi-Dirac statistics.

1. Bose-Einstein statistics. This is applicable to the identical, indistinguishable particles of zero or integral spin. These particles are called *bosons*. The examples of bosons are helium atoms at low temperature and the photons.

2. Fermi-Dirac statistics. This is applicable to the identical, indistinguishable particles of half-spin. These particles obey Pauli exclusion principle and are called *fermions*. The examples of fermions are electrons, protons, neutrons etc. In such systems of particles, no more than one particle can be in one quantum state.

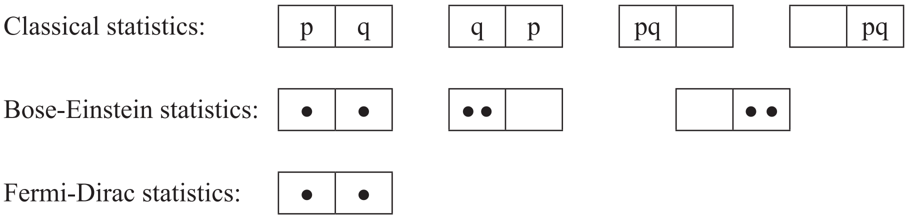
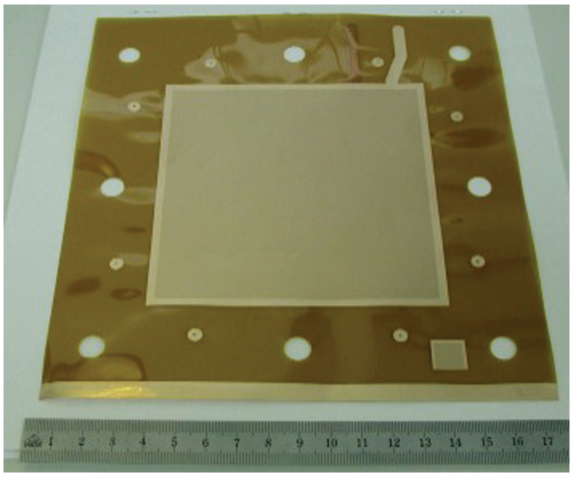


Fig. 76.3

The essential difference between the three statistics may be illustrated in the following simple manner. Let there be only two particles of a collection and only two cells to be occupied. Then Fig. 76.3 illustrates the essential difference between the three statistics.

76.9 Electron Gas



Electron Gas Multiplier.

Various properties of the metals such as electrical and thermal conductivities can be explained on the assumption that the electrons in the metal are free to move exactly like the particles of a gas. We can study the behaviour of the electrons on the basis of F-D statistics by considering them to form an electron gas in the metal.

Expression for Fermi Energy

The "free electron gas" in a solid obeys Fermi-Dirac statistics. Suppose in an assemblage of fermions, there are $M(E)$ allowed quantum states in an energy range between E and $E + dE$ and $N(E)$ is the number of particles in the same range. Then $N(E)$ quantum states are filled and $M(E) - N(E)$ are vacant. The F-D distribution function $f(E)$ is defined as :

$$f(E) = \frac{N(E)}{M(E)} = \frac{1}{1+e^{(E-E_f)/kT}} \quad ...(1)$$

$N(E) / M(E)$ represents the fraction of the possible quantum states which is occupied. The distribution of electrons among the levels is usually described by the *distribution function* $f(E)$. It is defined as the probability that the level E is occupied by an electron. Thus, if the level is certainly empty, then $f(E) = 0$. If the level is certainly full, then $f(E) = 1$. In general $f(E)$ has a value between zero and unity.

The distribution function for electrons at $T = 0\ K$ has the form

$$f(E) = 1 \text{ when } E < E_f$$

and $$f(E) = 0 \text{ when } E > E_f \qquad \text{...(2)}$$

That is, all levels below E_f are completely filled and all levels above E_f are completely empty. This function is plotted in Fig. 76.4 which shows the discontinuity at the Fermi-energy. As the temperature rises, $f(E)$ changes from 1 to 0 more and more gradually as shown in Fig. 76.4.

For $E = E_f$, $f(E) = \dfrac{1}{1+e^0} = \dfrac{1}{2}$, at all temperatures.

Thus, the probability of finding an electron with energy equal to the Fermi-energy in a metal is 1/2 *at any temperature.*

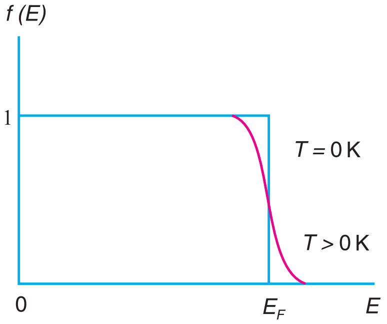


Fig. 76.4

Since the electrons are confined inside the crystal, their wave properties will limit the energy values which they may have. Let $g(E)\,dE$ be the number of quantum states available to electrons with energies between E and $E + dE$. It can be shown that

$$g(E)\,dE = \frac{8\sqrt{2}\,\pi V m^{3/2}}{h^3}\sqrt{3}\,dE \qquad \text{...(3)}$$

where m is the mass of the electron and V is the volume of the electron gas.

We can calculate the Fermi-energy E_F by filling up the energy states in the metal sample with the N free electrons it contains in order of increasing energy, starting from $E = 0$. The highest state to be filled will then have the energy $E = E_F$ by definition. The number of electrons that can have the same energy E is equal to the number of states that have this energy, since each state is limited to one electron. Hence,

$$N = \int_0^{E_f} g(E)\,dE = \frac{8\sqrt{2}\,\pi V m^{3/2}}{h^3}\int_0^{E_F} E^{1/2}\,dE$$

$$= \frac{16\sqrt{2}\,\pi V m^{3/2}}{3h^3}E_F^{3/2}$$

$$\therefore \qquad E_F = \left(\frac{h^2}{2m}\right)\left(\frac{3N}{8\pi V}\right)^{2/3} \qquad \text{...(4)}$$

The quantity N/V is the density of free electrons.

N/V represents the number of free electrons per unit volume of the metal.

An effective temperature of the electron-gas, known as the *Fermi-temperature*, is defined by

$$T_F = E_F / k$$

EXAMPLE 1. *An electron gas obeys the Maxwell-Boltzmann statistics. Calculate the average thermal energy (in eV) of an electron in the system at* 300 *K.*

SOL.

$$E = \frac{3}{2}kT = \frac{3}{2}\times(1.38\times10^{-23})\times 300 \text{ J}$$

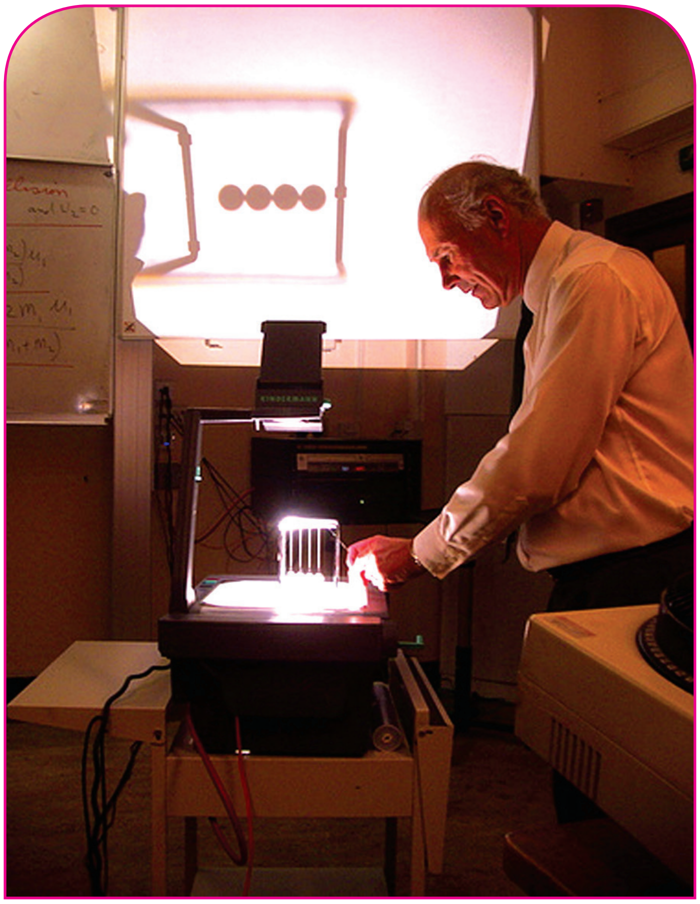
$$= \frac{3\times(1.38\times10^{-23})\times300}{2\times(1.6\times10^{-19})}\text{ eV} = 0.039 \text{ eV}\cdot$$

BLANK

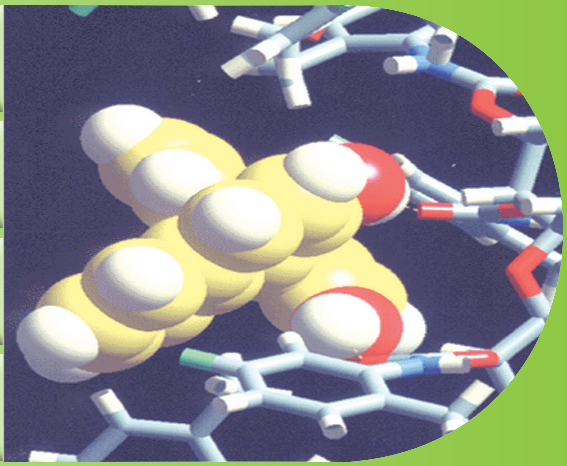
PART–X

MATERIAL SCIENCE

77. Material Science



CHAPTER 77



MATERIAL SCIENCE

AT A GLANCE

77.1	The Seven Crystal Systems	**77.2**	Absence of Five-fold Symmetry Axis
77.3	Definitions	**77.4**	Simple Cubic (SC) Structure
77.5	Body Centred Cubic (BCC) Structure	**77.6**	Face Centred Cubic (FCC) Structure
77.7	Calculation of Lattice Constant	**77.8**	Hexagonal Close Packed (HCP) Structure
77.9	Diamond Cubic Structure	**77.10**	Summary
77.11	Defects in Crystals	**77.12**	Classification of Crystal Imperfections (or Defects)
77.13	Point Defects	**77.14**	To Derive an Expression for the Number of Schottky Defects at a Given Temperature
77.15	To Derive an Expression for the Number of Frenkel Defects at a Given Temperature	**77.16**	To Derive an Expression for the Number of Vacancies at a Given Temperature
77.17	Line Defects	**77.18**	Surface Defects
77.19	Effects of Crystal Imperfection	**77.20**	Ferrites and their Applications
77.21	Ferromagnetic Domains	**77.22**	Four Types of Energy Involved in the Process of Domain Growth
77.23	Reversible and Irreversible Domains	**77.24**	Explanation of Hysteresis Curve on the Basis of Domain Theory
77.25	Crystal Growth	**77.26**	Crystal Growth from Solution
77.27	Crystal Growth from Melt-Czochralski Method	**77.28**	Crystal Growth from Vapour Phase
77.29	Weiss Theory of Paramagnetism	**77.30**	Heisenberg's Interpretation of Internal Field or Weiss Molecular Field
77.31	Quantum Theory of Ferromagnetism	**77.32**	Fundamental Definitions in Dieletrics
77.33	Different Types of Electric Polarisation	**77.34**	Total Polarisation
77.35	Effect of Temperature on Polarisation	**77.36**	Dielectric loss
77.37	Determination of Dielectric Constant of a Dielectric Material	**77.38**	Dielectric Breakdown
77.39	Properties and Applications of Different Types of Insulating Materials		

77.1 The Seven Crystal Systems

(*i*) Cubic crystal. In this crystal (Fig. 77.1), all the three lengths of the unit cell are equal and are at right angles to each other.

$$a = b = c$$
$$\alpha = \beta = \gamma = 90°$$

The examples are NaCl, CaF_2.

(*ii*) Tetragonal crystal. In this crystal (Fig. 77.2), the two lengths of the unit cell are equal while the third one is longer. The three axes are perpendicular, *i.e.* ($a = b \neq c$ and $\alpha = \beta = \gamma = 90°$). The examples are $NiSO_4$, SnO_2, etc.

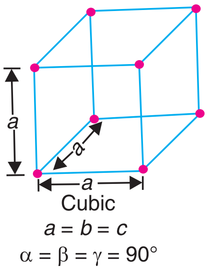

a
a
a
Cubic
$a = b = c$
$\alpha = \beta = \gamma = 90°$

Fig. 77.1

Perfect Cubic Crystal.

(*iii*) Orthorhombic crystal. In this crystal (Fig. 77.3), the lengths of unit cell are different but three axes are perpendicular to each other *i.e.*, ($a \neq b \neq c$ and $\alpha = \beta = \gamma = 90°$).

The examples are KNO_3, $BaSO_4$, etc.

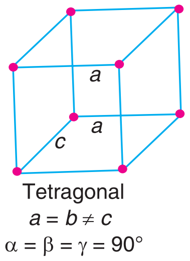

a
a
c
Tetragonal
$a = b \neq c$
$\alpha = \beta = \gamma = 90°$

Fig. 77.2

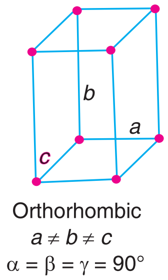

b
a
c
Orthorhombic
$a \neq b \neq c$
$\alpha = \beta = \gamma = 90°$

Fig. 77.3

(*iv*) Monoclinic crystal. In this crystal (Fig. 77.4), the lengths of unit cell are different. Two axes are right angles and the third is obliquely inclined *i.e.*, ($a \neq b \neq c$ and $\alpha = \beta = 90° \neq \gamma$).

The examples are Na_2SO_4, $FeSO_4$ etc.

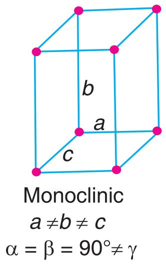

b
a
c
Monoclinic
$a \neq b \neq c$
$\alpha = \beta = 90° \neq \gamma$

Fig. 77.4

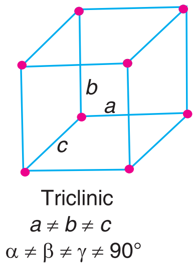

b
a
c
Triclinic
$a \neq b \neq c$
$\alpha \neq \beta \neq \gamma \neq 90°$

Fig. 77.5

(*v*) Triclinic crystal. In this crystal (Fig. 77.5), the lengths of unit cell are different and oblique to each other, *i.e.*, ($a \neq b \neq c$ and $\alpha \neq \beta \neq \gamma$).

The examples are $CuSO_4$, $K_2Cr_2O_7$, etc.

(*vi*) Trigonal (Rhombohedral) crystal. In this crystal (Fig. 77.6), all the three lengths of the unit cell are equal in length. Moreover, they are equally inclined to each other at an angle other than 90°.

$$a = b = c \text{ and } \alpha = \beta = \gamma \neq 90° .$$

The examples are $CaSO_4$, Calcite, As, Sb, Bi, etc.

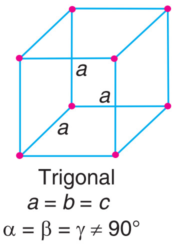


Fig. 77.6

(*vii*) Hexagonal crystal. In this crystal (Fig. 77.7), the three axes of unit cell are equal in one plane at 120° with each other fourth axis is perpendicular to this plane. *i.e.*, ($a = b \neq c$ and $\alpha = \beta = 90°$. $\gamma = 120°$).

The examples are SiO_2, AgI, Quartz, Zn, Cd, etc.

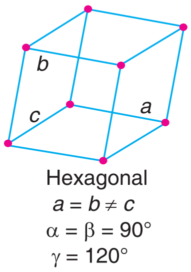
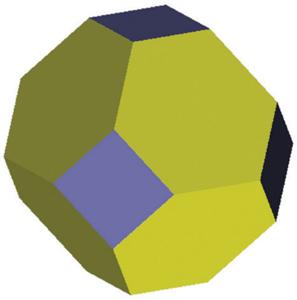


Fig. 77.7



Hexagonal Crystal.

77.2 Absence of Five-fold Symmetry Axis

Allowed rotation : An axis of *rotational symmetry* is an axis such that the lattice takes up an identical position on rotation through a suitable angle about it.

The angle of rotation must be of the form $2\pi/n$.

Here n is an integer and such an axis is called an n-fold axis.

A crystal can have only one, two, three, four, and six-fold axes of symmetry.

Proof. Let A (Fig. 77.8) be one of the points of a crystal lattice through which an axis of symmetry passes (perpendicular to the plane of the diagram). a is the lattice parameter. B is another point of the lattice separated by one lattice vector from A. Another identical axis of symmetry must pass through B. Now, rotate the crystal through an angle $\theta = 2\pi/n$, about the axis through A. Then, the point B, as well as the axis through B, takes the position B' . Similarly, a rotation about B takes A to A'.

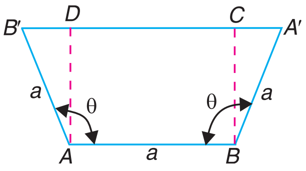


Fig. 77.8

By assumption A and B are lattice points of the same crystal and hence can replace each other by a parallel displacement. So it is apparent that $A'B'$ must also be a translational period of the lattice. Thus, if a is the shortest period in the given direction, then the distance $A'B' = pa$.

Here, p is an integer.

$$\begin{aligned} A'B' &= A'C + CD + DB' \\ &= a\sin\left(\theta - \frac{\pi}{2}\right) + a + a\sin\left(\theta - \frac{\pi}{2}\right) \\ &= a - 2a\cos\theta = pa \end{aligned}$$

i.e., $\cos\theta = (1 - p)/2.$

Since $|\cos\theta| \leq 1$, p can only have the values $p = 0, \pm 1, 2, 3$.

Now we have to find the values of θ and n.

p	$\cos\theta$	θ	$n = \frac{360}{\theta}$
0	$\frac{1}{2}$	60	6
+1	0	90	4
–1	+1	360 or 0	1
2	$-\frac{1}{2}$	120	3
3	–1	180	2

This proves that a crystal can have only 1, 2, 3, 4, and 6-fold axes of rotation and it cannot have a five fold rotation axis.

This is illustrated in Fig. 77.9. We observe that an array of pentagons cannot cover all space.

The array has point symmetry, *i.e.*, 5-fold rotational axis but not the translational symmetry.

Calculation of Number of Atoms per unit cell, Atomic Radius, Coordination Number and Packing Factor for Simple Cubic, BCC, FCC, HCP and Diamond structures.

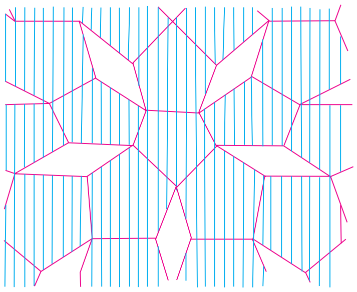


Fig. 77.9

77.3 Definitions

1. Atomic radius r : Atomic radius is defined as half the distance between two nearest neighbouring atoms in a crystal.

2. Nearest Neighbour Distance (2r): The distance between the centres of two nearest neighbouring atoms is called *nearest neighbour distance*. If r is the radius of the atom, nearest neighbour distance is $2r$.

3. Co-ordination number : *Coordination number is the number of nearest neighbour atoms that an atom has in the given crystal structure.*

Explanation. Each atom in a unit cell has some nearest neighbouring atoms. The number of these nearest neighbours to a given atom in crystal lattice is called the coordination number (CN) in crystal lattice. Thus the number of atoms in the immediate neighbourhood of a given atom in a crystal is called the co-ordination number (CN).

The coordination number is calculated by taking the nearest neighbouring atoms at a distance, $2r$.

It gives an indication of the closeness of the packing of the atoms.

4. Packing factor : *Atomic packing factor is the ratio of the volume of atoms in the unit cell to the volume of the unit cell.*

Explanation. The ratio of volume occupied by atoms in a unit cell to the total volume of the unit cell is called *packing factor*.

Packing Factor, P.F. $= v/V$.

Here, v = volume occupied by the atoms in the unit cell and

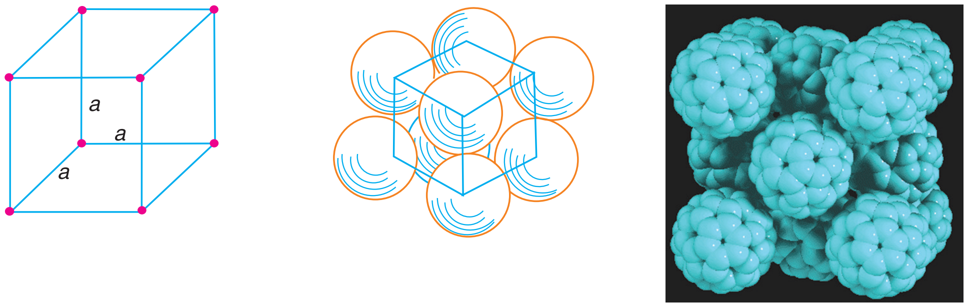
V = total volume of the unit cell.

A high value of P.F. indicates that the atoms are very closely packed in the unit cell and there is very small unoccupied space. A low value of P.F. indicates that the atoms are very loosely packed in the unit cell and there is enormous unoccupied space.

We assume that the atoms are hard spheres and that the nearest neighbours touch each other.

77.4 Simple Cubic (SC) Structure

The cubic unit cell for simple cubic lattice is shown in Fig. 77.10.



(*a*) atomic-site model (*b*) hard sphere model ***Simple Cubic Structure.***

Fig. 77.10

(*i*) Number of atoms per unit cell

There are 8 atoms at the 8 corners of the cell.

Each corner atom is shared by 8 unit cells that adjoin at each corner.

The share of each corner atom to a unit cell is 1/8 of an atom.

Total number of atoms per unit cell is $8\times\frac{1}{8}=1$.

(*ii*) Atomic radius , *r*

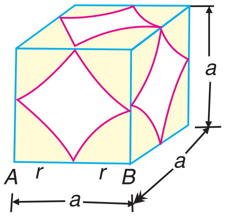


Fig. 77.11

In a *SC* cell the atoms are in contact along the edges of the cube (Fig. 77.11).

a is length of side of cubic cell.

The atoms touch along cube edges.

Nearest neighbour distance,

$$2r = a$$

$$\therefore \quad r = a/2.$$

(*iii*) Coordination number, CN

Each corner atom touches four atoms in its horizontal plane, and two atoms, one vertically above it and one vertically below it. So there are six equally spaced nearest neighbour atoms each at a distance '*a*' from that atom. Therefore, the coordination numbr is six.

$$CN = 6.$$

(*iv*) Packing Factor

$$\text{Packing factor} = \frac{\text{Volume of all atoms in the unit cell}}{\text{Volume of the unit cell}}$$

$$= \frac{\text{No. of atoms in unit cell} \times \text{Volume of one atom}}{\text{Volume of the unit cell}}$$

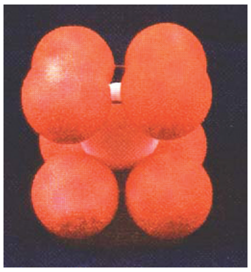
$$= \frac{1\times\frac{4}{3}\pi r^3}{(a)^3} = \frac{4\pi r^3}{3(2r)^3} \qquad (\because a = 2r)$$

$$= \frac{\pi}{6} = 0.52\,.$$

$\therefore$ Packing factor = 52%

Simple cubic structure is a loosely packed structure.

Polonium is the only element which exhibits this structure.



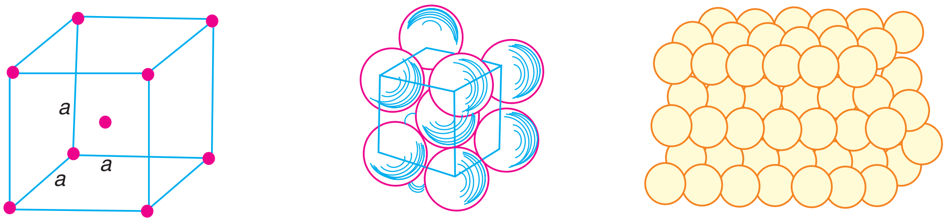
Body Centred Cubic Structure.

77.5 Body Centred Cubic (BCC) Structure

Fig. 77.12 (*a*) shows the arrangement of lattice points in a *BCC* cell.

Fig. 77.12 (*b*) shows the arrangment of atoms represented as hard spheres.

Fig. 77.12 (c) shows the BCC repeating structure.



(*a*) atomic site model (*b*) hard sphere model (*c*) repeating structure in a crystal

Fig. 77.12

Let us now compute the characteristics of a *BCC* cell.

(*i*) Number of atoms per unit cell

There are 8 atoms at the 8 corners, each shared by 8 unit cells and one atom at the body centre. The atom at the centre of the body of the cell wholly belongs to the unit cell.

Thus, the total number of atoms per unit cell is $8 \times \frac{1}{8} + 1 = 2$.

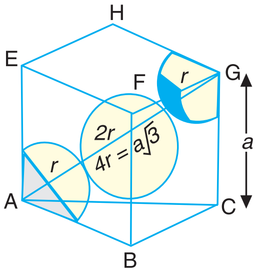


Fig. 77.13

(*ii*) Atomic radius, r

In this case, atoms touch each other along the diagonal of the cube (Fig. 77.13).

The length of the body diagonal = 4*r*.

$$AG^2 = AC^2 + CG^2 = (AB^2 + BC^2) + CG^2$$

$$(4r)^2 = 3a^2$$

Atomic radius $r = \frac{\sqrt{3}\,a}{4}$

Lattice constant $a = \frac{4}{\sqrt{3}} r\,\cdot$

(*iii*) Coordination number, *CN*

The unit cell contains one atom at each of the eight corners of the unit cell and one atom at the body centre. The nearest atom to the corner atom is the body centred atom and not the other corner atoms. The number of nearest neighbours is eight (Fig. 77.14).

The body centre atom is in contact with all the eight corner atoms. Therefore, the coordination number is eight.

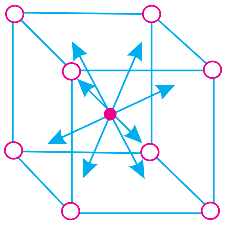


Fig. 77.14

$$CN = 8.$$

(*iv*) Packing Factor

$$\text{Packing factor} = \frac{\text{Volume of all atoms in the unit cell}}{\text{Volume of the unit cell}}$$

$$= \frac{\text{No. of atoms in unit cell} \times \text{Volume of one atom}}{\text{Volume of the unit cell}}$$

$$= \frac{2\times(4/3)\pi r^3}{a^3} = \frac{8\pi r^3}{3a^3}$$

$$= \frac{\left(\frac{\sqrt{\ }}{\ }\right)}{\ } \qquad \left(\because r = \frac{\sqrt{3}a}{4}\right)$$

$$= \frac{\sqrt{3}\pi}{8} = 0.68\,.$$

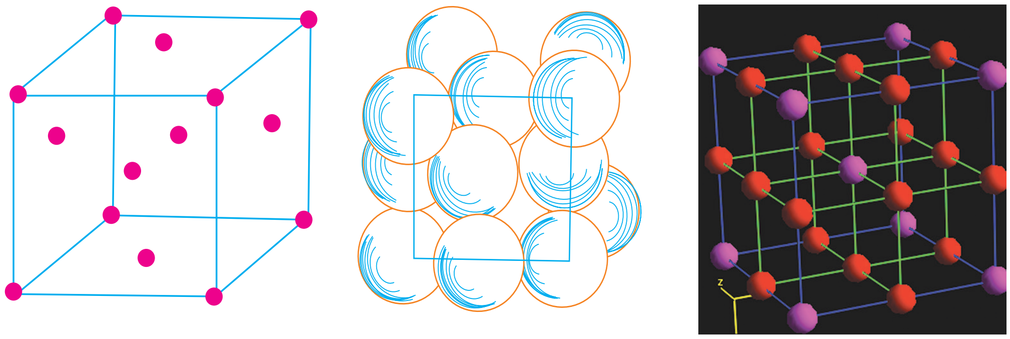
∴ Packing factor = 68%.

Thus the packing factor is equal to 68%. So it is not a closely packed structure.

Tungsten, sodium, iron and chromium have this type of structure.

77.6 Face Centred Cubic (FCC) Structure

The unit cell of this structure is shown in Fig. 77.15.



(*a*) atomic site model (*b*) hard sphere model *Face Centred Cubic Structure.*

Fig. 77.15

The face centred cubic unit cell is a nonprimitive cell.

It has six atoms at the centres of six faces and eight atoms at the eight corners of the cube.

We shall now determine the characteristics of the *FCC* cell.

(*i*) Number of atoms per unit cell

There are 8 atoms, each shared by 8 unit cells at 8 corners and 6 face-centred atoms, each shared by 2 cells.

The total number of atoms per unit cell = contribution due to 8 corner atoms + contribution due to 6 face centred atoms

$$= \left(\frac{1}{8}\times 8\right) + \left(\frac{1}{2}\times 6\right) = 1 + 3 = 4$$

Thus the total number of atoms per unit cell is 4.

(*ii*) Atomic radius, *r*

In a *FCC* cell, atoms are in contact along the face diagonal of the cube (Fig. 77.16).

Length of the face diagonal, $AF = 4r$.

$$AF^2 = AB^2 + BF^2$$
$$(4r)^2 = a^2 + a^2 = 2a^2$$
$$4r = \sqrt{2}a$$

$\therefore$ Atomic radius, $$r = \frac{a}{2\sqrt{2}}$$

Lattice constant $a = 2\sqrt{2}\,r$.

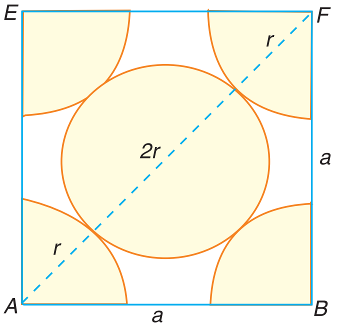


Fig. 77.16

(*iii*) Coordination number CN

In this case, the nearest neighbours of any corner atom are the face centred atoms of the surrounding unit cells. Any corner atom has four such atoms in its own plane, four in a plane above it and four in a plane below it.

Therefore, the coordination number $CN = 12$.

Theoretically, a sphere can touch simultaneously 12 identical spheres at the maximum in a three dimensional structure. Thus, an *FCC* cell has the maximum value for the coordination number.

(*iv*) Packing Factor

$$\text{Packing Factor} = \frac{\text{Volume of all atoms in the unit cell}}{\text{Volume of the unit cell}}$$
$$= \frac{\text{No. of atoms in unit cell} \times \text{Volume of one atom}}{\text{Volume of unit cell}}$$
$$= \frac{4\times\frac{4}{3}\pi r^3}{a^3}$$
$$= \frac{16}{3}\cdot\frac{\pi r^3}{\left(2\sqrt{2}r\right)^3} \qquad (\because a = 2\sqrt{2}r)$$
$$= \frac{\pi}{3\sqrt{2}} = 0.74$$

$\therefore$ Packing factor = 74%

When compared with *SC* and *BCC*, this has the highest packing factor.

Most of the metals like copper, aluminium, lead and silver have this structure.

77.7 Calculation of Lattice Constant

Consider a cubic lattice of lattice constant *a*.

Let ρ be the density of the crystal.

Volume of the unit cell = a^3

$\therefore$ Mass in each unit cell = $\rho\, a^3$ ...(1)

Let n = No. of molecules (lattice points) per unit cell,

M = molecular weight and

N = Avogadro number (*i.e.*, number of molecules per kg mole of the substance).

$$\text{Mass of each molecule} = \frac{M}{N}$$

$$\text{Mass of each unit cell} = n \times \frac{M}{N} \qquad \text{...(2)}$$

From Eqs. (1) and (2), we have

$$\rho\, a^3 = n \,.\, \frac{M}{N} \text{ or } a^3 = \frac{nM}{\rho N}$$

$$\therefore \qquad a = \left(\frac{nM}{N\rho}\right)^{1/3}. \qquad \text{... (3)}$$

EXAMPLE 1. *A substance with face centred cubic lattice has density 6250 kg/m^3 and molecular weight 60.2. Calculate the lattice constant a. Given: Avogadro number = 6.02 × 10^{26} (kg.mole)$^{-1}$.*

SOLUTION.
$$a = \left(\frac{nM}{N\rho}\right)^{1/3}$$

Here n = number of molecules per unit cell in FCC lattice = 4

$$M = 60.2,\ N = 6.02 \times 10^{26} \text{ and } \rho = 6250 \text{ kg/m}^3.$$

$$a = \left[\frac{4\times 60.2}{\left(6.02\times 10^{26}\right)\times 6250}\right]^{\frac{1}{3}} = 0.4\times 10^{-9}\,m = 0.4\,nm\,\cdot$$

EXAMPLE 2. *NaCl crystals have FCC structure. The density of sodium chloride is 2180 kg m^{-3}. Calculate the distance between two adjacent atoms. Atomic weight of sodium = 23 and that of chlorine = 35.5.*

SOLUTION.
$$a = \left(\frac{nM}{N\rho}\right)^{1/3}$$

Here $n = 4$, $M = 23 + 35.5 = 58.5$, $N = 6.02 \times 10^{26}$ $(kg\ mole)^{-1}$ and $\rho = 2180$ kg m^{-3}.

$$\therefore \qquad a = \left[\frac{4\times 58.5}{\left(6.02\times 10^{26}\right)\times 2180}\right]^{\frac{1}{3}} = 0.5657\times 10^{-9}\,m\,\cdot$$

The distance between two adjacent atoms is half of this value, *i.e.*,

$$d = \frac{a}{2} = \frac{0.5657\times 10^{-9}}{2}\text{m} = 0.2828\times 10^{-9}\text{ m} = 0.2828\ \text{ nm.}$$

EXAMPLE 3. *Copper has FCC structure and the atomic radius is 1.278 Å. Calculate its density.*

SOLUTION.
$$\rho = \frac{nM}{Na^3}$$

Here n = 4 atoms, $M = 63.54$, $N = 6.023 \times 10^{26}$ (kg mole)$^{-1}$, $r = 1.278$ Å $= 1.278 \times 10^{-10}$ m.

$$a = 2\sqrt{2}\,r = 2\times 1.414\times(1.278\times 10^{-10})\,\text{m} = 3.614\times 10^{-10}\,\text{m.}$$

$$\rho = \frac{4\times 63.54}{\left(6.023\times 10^{26}\right)\times\left(3.614\times 10^{-10}\right)^3} = 8940\,\text{kg}\,\text{m}^{-3}.$$

EXAMPLE 4. *Lead is a face centred cubic with an atomic radius of 1.746 Å. Find the spacing of (i) 200 planes and (ii) 220 planes.*

SOLUTION.
$$d_{hkl} = \frac{a}{\sqrt{h^2+k^2+l^2}}.$$

The lattice constant of unit cell of lead (FCC) structure is related to its atomic radius by the relation

$$a = \frac{4r}{\sqrt{2}} = \frac{4\times 1.746}{\sqrt{2}} = 4.93\,\text{Å}$$

(*i*) For (200) plane,

$$d_{200} = \frac{4.93}{\sqrt{(2)^2+(0)^2+(0)^2}} = \frac{4.93}{2} = 2.465\,\text{Å}.$$

(*ii*) For (220) planes

$$d_{200} = \frac{4.93}{\sqrt{(2)^2+(2)^2+(0)^2}} = \frac{4.93}{2\sqrt{2}} = 1.74\,\text{Å}.$$

EXAMPLE 5. *Copper has an FCC structure with lattice constant a = 3.61 Å. Calculate the radius of copper-atom.*

SOLUTION.
$$r = \frac{\sqrt{2}a}{4} = \frac{\sqrt{2}\times 3.61}{4}\,\text{Å} = 1.28\,\text{Å}.$$

EXAMPLE 6. *Metallic iron changes from BCC to FCC form at 910°C. At this temperature, the atomic radii of the iron atoms in the two structures are 1.258 Å and 1.292 Å respectively. Calculate the percent volume change during this structural change.*

SOLUTION.

(*i*) In BCC, $a = \frac{4r}{\sqrt{3}} = \frac{4\times 1.258\text{Å}}{\sqrt{3}} = 2.905\ \text{Å} = 2.905\times 10^{-10}\,\text{m}.$

Volume of unit cell = $a^3 = (2.905\times 10^{-10})^3 = 24.515\times 10^{-30}\,\text{m}^3$.

Number of atoms per unit cell = 2

$\therefore$ Volume occupied by one atom = $\frac{24.515\times 10^{-30}}{2} = 12.26\times 10^{-30}\,\text{m}^3$.

(*ii*) In FCC, $a = 2\sqrt{2}r = 2\sqrt{2}\times 1.292\,\text{Å}$

$$= 2\sqrt{2}\times(1.292\times 10^{-10})\,\text{m} = 3.654\times 10^{-10}\,\text{m}.$$

Volume of unit cell = $a^3 = (3.654\times 10^{-10})^3 = 48.787\times 10^{-30}\,\text{m}^3$.

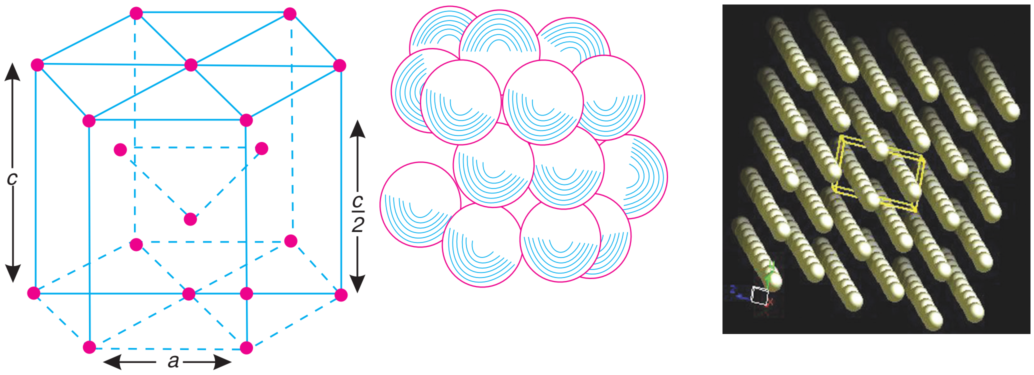
Number of atoms per unit cell = 4

Volume occupied by one atom = $\frac{48.787\times 10^{-30}}{4} = 12.2\times 10^{-30}\,\text{m}^3$

$\therefore$ Volume change in % $= \frac{12.26-12.2}{12.26}\times 100 = 0.489\cdot$

77.8 Hexagonal Close Packed (HCP) Structure

The conventional unit cell for hexagonal close packed structure is shown in Fig. 77.17.



(*a*) atomic site model (*b*) hard sphere model *Hexagonal Close Packed (HCP) Structure.*

Fig. 77.17

There are three layers of atoms in the unit cell.

(*i*) At the bottom layer, central atom has six nearest neighbour atoms in the same plane.

(*ii*) Top layer has the same atomic arrangement as the bottom layer.

(*iii*) At $c/2$ distance from the bottom layer there is a middle layer containing three atoms.

Let us now compute the characteristics of *HCP* cell.

(*i*) Number of atoms per unit cell

The top layer contains 6 atoms at the corners and one atom at the centre.

The corner atoms are shared between 6 unit cells. Hence 6 atoms contribute only 1 atom to the unit cell.

The central atom is shared between 2 unit cells and therefore contributes $\frac{1}{2}$ atom to unit cell.

Hence the top layer contributes $1+\frac{1}{2}=\frac{3}{2}$ atoms to the unit cell.

Similarly, the bottom layer contributes $\frac{3}{2}$ atoms to the unit cell.

The three atoms of the middle layer lie within the volume of the unit cell and hence fully (3 atoms) contribute to the cell.

Thus, the total number of atoms/unit cell.

$$= \frac{3}{2}+\frac{3}{2}+3=6\,.$$

(*ii*) Atomic radius, r

The atoms are in contact along the edges of the hexagon (Fig. 77.18)

$$\therefore \qquad 2r = a$$

$$\text{or} \qquad r = a/2.$$

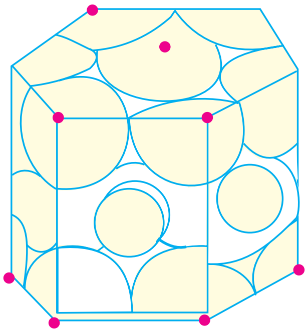


Fig. 77.18

(*iii*) Coordination number , CN

Consider the central atom at the top layer of HCP structure. It touches 6 atoms in its plane, 3 atoms in the plane just below it and 3 atoms in the plane just above it. Hence the coordination number is 12.

(*iv*) Volume of the unit cell

The volume of the unit cell is determined by computing the area of the base of the unit cell and then by multiplying it by the cell height. The area of the base of the unit cell is the area of the hexagon *ABDEFG* (Fig. 77.19). It is equal to the sum of the areas of the six equal-sized triangles.

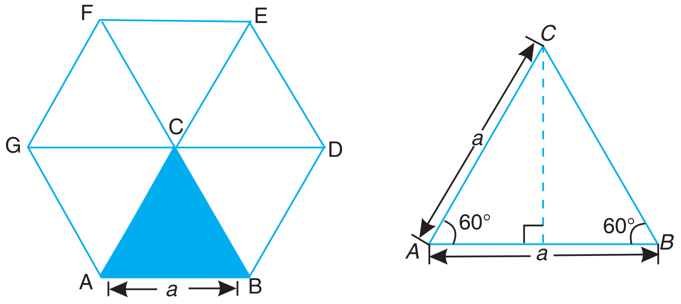


Fig. 77.19

$$\therefore \text{ Area of hexagon } ABDEFG = 6 \text{ (area of } \Delta \text{ ABC)}$$

$$= 6\times\left[\frac{1}{2}\times a\times(a\sin 60^\circ)\right]$$

$$= 3a^2\left(\frac{\sqrt{3}}{2}\right) = \frac{3\sqrt{3}}{2}a^2.$$

$$\therefore \text{ Volume of the unit cell, V} = \text{(area of the base)} \times \text{(Height of the cell)}$$

$$= \left(\frac{3\sqrt{3}}{2}a^2\right)c\,\cdot$$

Calculation of c/a ratio

Let c be the height of the unit cell and a be its edge. Also $a = 2r$.

The three body atoms lie in a horizontal plane at a height $c/2$ from the orthocentres of alternate equilateral triangles in the base or at top of the hexagonal cell.

From Fig. 77.20,

$$(2r)^2 = x^2 + \left(\frac{c}{2}\right)^2$$

But, $$x = \frac{2}{3}AN = \frac{2}{3}\left(a\frac{\sqrt{3}}{2}\right) = \frac{a}{\sqrt{3}}$$

Also $$2r = a$$

$\therefore$ $$a^2 = \left(\frac{a}{\sqrt{3}}\right)^2 + \left(\frac{c}{2}\right)^2 = \frac{a^2}{3} + \frac{c^2}{4}$$

$$\frac{c^2}{4} = a^2 - \frac{a^2}{3} = \frac{2a^2}{3}$$

or $$\frac{c^2}{a^2} = \frac{8}{3}$$

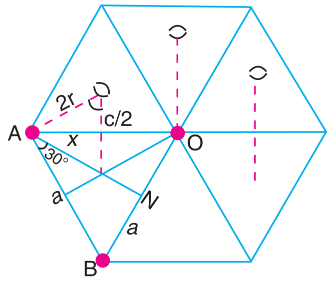


Fig. 77.20

$$\therefore \qquad \frac{c}{a} = \sqrt{\frac{8}{3}} = 1.633\cdot$$

(*v*) Calculation of packing factor

$$\text{Packing factor} = \frac{\text{Volume of all atoms in the unit cell}}{\text{Volume of the unit cell}}$$

$$= \frac{\text{No. of atoms in unit cell} \times \text{Volume of one atom}}{\text{Volume of the unit cell}}$$

$$= \frac{6\times\frac{4}{3}\pi r^3}{\frac{3\sqrt{3}}{2}a^2c} = \frac{16\pi r^3}{3\sqrt{3}\,a^2c}$$

$$= \frac{16\pi(a/2)^3}{3\sqrt{3}\,a^2c} \qquad \left(\because r = \frac{a}{2}\right)$$

$$= \frac{2\pi a^3}{3\sqrt{3}\,a^2c} = \frac{2\pi}{3\sqrt{3}}\frac{a}{c}$$

$$= \frac{2\pi}{3\sqrt{3}}\sqrt{\frac{3}{8}} \qquad \left(\because \frac{c}{a} = \sqrt{\frac{8}{3}}\right)$$

$$= \frac{\pi}{3\sqrt{2}} = 0.74$$

∴ Packing factor = 74%

Since the density of packing is 74%, it is a close packed structure.

Magnesium, zinc and cadmium crystallize in this structure.

EXAMPLE 1. *Zinc has HCP structure. The height of the unit cell is 0.494 nm. The nearest neighbour distance is 0.27 nm. The atomic weight of zinc is 65.37. Calculate the volume of the unit cell and density of zinc.*

SOLUTION. Here, $a = 2r = 0.27 \times 10^{-9}$ m; $c = 0.494 \times 10^{-9}$ m.

Volume of the unit cell.

$$V = \frac{3\sqrt{3}\,a^2c}{2} = \frac{3\times\sqrt{3}\times(0.27\times10^{-9})^2\times(0.494\times10^{-9})}{2}$$

$$= 9.356 \times 10^{-29}\ \text{m}^3.$$

Let ρ = density of the crystal.

Mass of the unit cell = $\rho V = (9.356 \times 10^{-29})\rho$ kg.

The number of atoms belonging to a unit cell of HCP = 6.

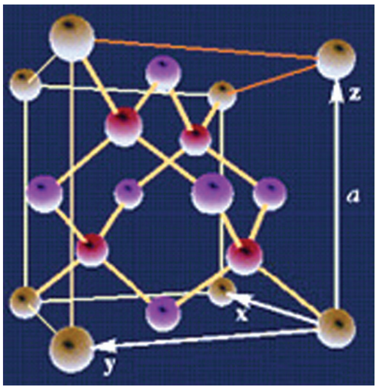
If M is the atomic weight and N the Avogadro number, then

$$\text{mass per unit cell} = \frac{6M}{N} = \frac{6\times65.37}{6.02\times10^{26}}$$

$$(9.356 \times 10^{-29})\ \rho = \frac{6\times65.37}{6.02\times10^{26}}$$

$$\rho = \frac{6\times65.37}{6.02\times10^{26}\times(9.356\times10^{-29})}$$

$$\rho = 6964\ \text{kg m}^{-3}.$$

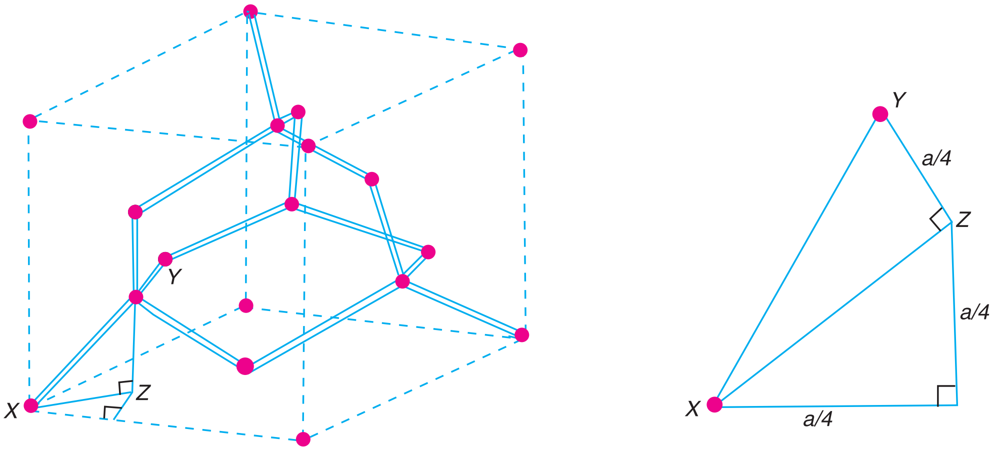


Diamond Cubic Structure.

77.9 Diamond Cubic Structure

The diamond structure is a combination of two inter-penetrating face-centred cubic (FCC) sublattices.

The unit cell is sketched in Fig. 77.21 (*a*).



(a) (b)

Fig. 77.21

One sublattice, say *X*, **has its origin at (0, 0, 0)**. The other sublattice, *Y*, has its origin quarter of the way along the body diagonal *i.e.,* at the point (*a*/4, *a*/4, *a*/4).

The diamond cubic structure is loosely packed, since each atom has only four nearest neighbours.

(*i*) Number of atoms per unit cell

In the unit cell, in addition to the eight corner atoms, there are six face centred atoms and four more atoms are located inside the unit cell. Each corner atom is shared by eight adjacent unit cells and each face centred atom is shared by two unit cells. Hence the total number of atoms per unit cell is,

$$\left(\frac{1}{8}\times 8\right)+\left(\frac{1}{2}\times 6\right)+4 \;=\; 8\cdot$$

(*ii*) Atomic radius, r

From Fig. 77.21 (*b*),

$$XZ^2 = \left(\frac{a}{4}\right)^2+\left(\frac{a}{4}\right)^2=\frac{a^2}{8}$$

Also $XY^2 = XZ^2 + YZ^2 = \frac{a^2}{8} + \frac{a^2}{16} = \frac{3a^2}{16}$.

But $XY = 2r$

$$\therefore \qquad (2r)^2 = \frac{3a^2}{16}$$

$$r = \frac{\sqrt{3}}{8}a$$

Lattice constant $a = \dfrac{8r}{\sqrt{3}}$.

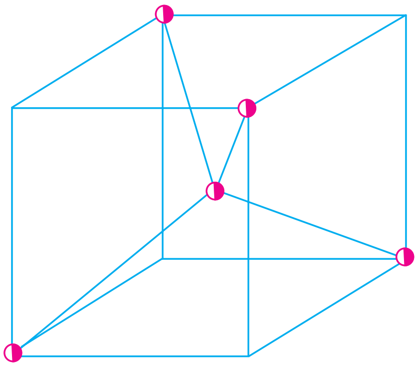


Fig. 77.22

(*iii*) Coordination Number

In a diamond crystal, the carbon atoms are linked by directional covalent bonds. Each carbon atom forms covalent bonds with four other carbon atoms that occupy four corners of a cube in a tetrahedral structure (Fig. 77.22).

In the diamond lattice, each atom has four nearest neighbours with which it forms covalent bonds. Thus the coordination number of diamond crystal is 4.

(*iv*) Packing Factor

$$\text{Packing factor} = \frac{\text{Volume of all atoms in the unit cell}}{\text{Volume of the unit cell}}$$

$$= \frac{\text{No. of atoms in unit cell} \times \text{Volume of one atom}}{\text{Volume of the unit cell}}$$

$$= \frac{8 \times \frac{4}{3}\pi r^3}{a^3}$$

$$= \frac{32\pi r^3}{3\left(\frac{8r}{\sqrt{3}}\right)^3} \qquad \left(\because a = \frac{8r}{\sqrt{3}}\right)$$

$$= \frac{\pi\sqrt{3}}{16} = 0.34$$

∴ Packing factor = 34 %

Thus it is a loosely packed structure.

Carbon, silicon, germanium and grey tin crystallize in diamond structure.

Example 1. *Silicon has the same structure as diamond and a bond length of 0.2351 nm. Estimate its density.*

Solution. In the case of diamond cubic structure,

Nearest neighbour distance $2r = \frac{\sqrt{3}}{4}a$.

Number of atoms per unit cell = 8

$$a = \frac{4}{\sqrt{3}}(2r) = \frac{4}{\sqrt{3}} \times 0.2351\,\text{nm} = 0.543\,\text{nm} = 0.543 \times 10^{-9}\,\text{m}$$

$$\rho = \frac{nM}{Na^3} = \frac{8 \times 28.09}{(6.02 \times 10^{26})(0.543 \times 10^{-9})^3} = 2331\,\text{kg m}^{-3}\,.$$

77.10 Summary

The characteristics of the unit cells of simple cubic, BCC, FCC, HCP and diamond cube structures **are summarized in Table 1**.

TABLE 1

Sl. No.	*Characteristic*	*(SC)*	*(BCC)*	*(FCC)*	*(HCP)*	*(DC)*
1.	Number of atoms per unit cell	1	2	4	6	8
2.	Atomic radius, r	$\frac{a}{2}$	$\frac{a\sqrt{3}}{4}$	$\frac{a}{2\sqrt{2}}$	$\frac{a}{2}$	$\frac{a\sqrt{3}}{8}$
3.	Nearest neighbour distance, $(2r)$	a	$\frac{a\sqrt{3}}{2}$	$\frac{a}{\sqrt{2}}$	a	$\frac{a\sqrt{3}}{4}$
4.	Coordination number	6	8	12	12	4
5.	Unit cell volume	a^3	a^3	a^3	$\frac{3\sqrt{3}}{2}a^2c$	a^3
6.	Packing factor	0.52	0.68	0.74	0.74	0.34
7.	Examples	Polonium	Iron, Barium, Chromium	Aluminium, Copper Lead, Gold	Magnesium, Zinc, Cadmium	Germanium, Silicon, Carbon diamond

DEFECTS IN CRYSTALS

77.11 Introduction

In an *ideal crystal*, the atomic arrangement is perfectly regular and continuous throughout. Thus an ideal crystal is a perfect one in all respects.

Real crystals usually contain several types of imperfections, such as substitutional and interstitial atoms, as well as vacancies or holes. Dislocations and surface defects are also usually present in crystals. The mechanical, electrical and magnetic properties of engineering crystalline solids, particularly metals and alloys, are profoundly affected by the imperfections in the crystals.

The properties of solids such as, strength of crystals, luminescence, colour of crystals, diffusion, crystal growth, plasticity etc., cannot be explained on the basis perfect periodicity concept. Such properties are greatly affected by the lattice defects or atomic imperfections arising due to irregularities of atomic arrays in crystals. These properties are called defect or structure sensitive properties.

The term imperfection or defect is used to describe any deviation from the perfect periodic array of atoms in the crystals.

77.12 Classification of Crystal Imperfections (or Defects)

Crystalline imperfections can be classified on the basis of their geometry under the following four main divisions.

1. POINT DEFECTS (zero dimensional defects)
   - (*a*) Vacancies
   - (*b*) Interstitialcies
   - (*c*) Impurities
   - (*d*) Electronic defects
2. LINE DEFECTS (one dimensional defects)
   - (*a*) Edge dislocation
   - (*b*) Screw dislocation

3. SURFACE DEFECTS (two dimensional defects)
   - (*a*) Grain boundaries
   - (*b*) Tilt boundaries
   - (*c*) Twin boundaries
   - (*d*) Stacking faults
   - (*e*) Ferromagnetic domain walls
4. VOLUME DEFECTS (three dimensional defects)
   - (*a*) Large voids
   - (*b*) Cracks

77.13 Point Defects

Point imperfections are also called zero dimensional imperfections. They are imperfect point-like regions in the crystal. One or two atomic diameters is the typical size of a point imperfection. In a crystal lattice, point defect is one which is completely local in its effect, *e.g.*, a vacant lattice site.

The introduction of point defect into a crystal increases its internal energy as compared to that of the perfect crystal.

(*i*) If the point defect is a vacancy, then there is no bonding. Hence the value of the mechanical strength at that point is reduced.

(*ii*) If an impurity atom is present instead of the original atom, in the lattice, there must be some strain due to the different size of the impurity atom.

*The number of defects at equilibrium at a certain temperature can be determined from the following equation :

$$n_d = Ne^{-Ed/kT}$$

Here, n_d is the number of defects,

N is the total number of atomic sites per cubic metre or per mole,

E_d is the energy of activation necessary to form the defect,

k is the Boltzmann constant, and

T is the absolute temperature.

77.13.1 Vacancies

An empty site of an atom in a crystal is called vacancy.

A vacancy implies an unoccupied atom position within a crystal lattice (Fig. 77.23).

The atoms surrounding a vacancy tend to be closer together, thereby distorting the lattice planes.

Vacancies may occur as a result of imperfect packing during the original crystallization or they may arise from **thermal vibrations** of atoms at elevated temperatures, because as thermal energy is increased there is a higher probability that individual atoms will jump out of their position of lowest energy.

It may be shown by thermodynamic reasoning that lattice vacancies are a stable feature of metals at all temperatures above absolute zero.

The everyday industrial processes of annealing, homogenization, precipitation, sintering, surface

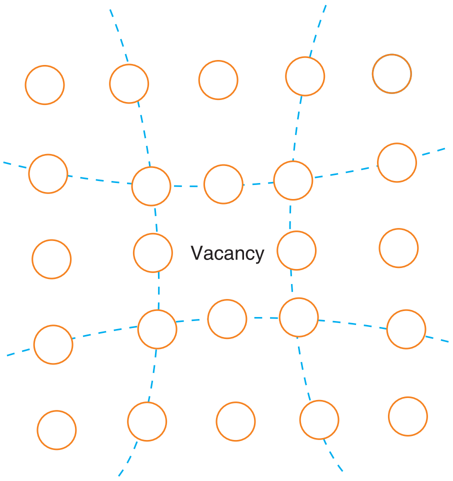


Fig. 77.23

hardening, oxidation and creep, all involve to varying degrees, the transport of atoms through the lattice with the help of vacancies.

Vacancies exist in a certain proportion in a crystal in thermal equilibrium, leading to an increase in the randomness of the structure.

Schottky defect. Schottky defect is closely related to vacancies. It is formed when an atom or an ion is removed from a normal lattice site and replaced in an average position on the surface of the crystal (Fig. 77.24). In other words, when vacancies are created by movements of atoms from positions inside the crystal to positions on the surface of the crystal, a *Schottky defect* is said to have been formed.

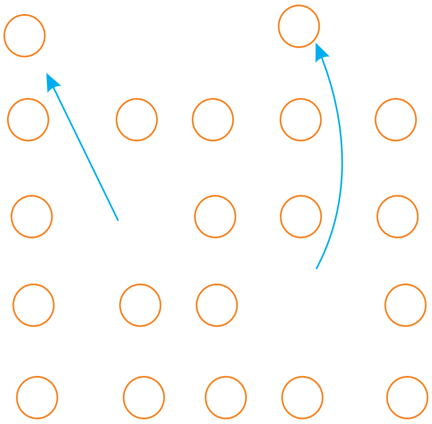


Fig. 77.24

Schottky defects in Ionic crystals

In ionic crystals, formation of point imperfections is subject to the requirement that the overall electrical neutriality is maintained. A pair of one cation and one anion can be missing from an ionic crystal as shown in Fig. 77.25.

When vacancies are created by movements of one anion and one cation from positions inside the crystal to positions on the surface of the crystal, a Schottky defect is said to have been formed.

The valency of the missing pair of ions should be equal to maintain the electrical neutrality. So these are normally generated in equal numbers of anion and cation vacancies in a crystal.

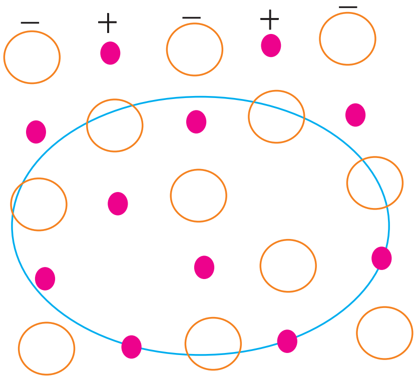


Fig. 77.25

Schottky defects are dominant in alkali halides.

77.13.2 Interstitialcies

An interstitial defect arises when an atom occupies a definite position in the lattice that is not normally occupied in the perfect crystal.

In interstitialcies, atoms occupy vacant positions between the atoms of the ideal crystal (Fig. 77.26).

The interstitial (atom) may be either a normal atom of the crystal or a foreign atom.

Interstitialcy produces atomic distortion because interstitial atom tends to push the surrounding atoms farther apart, unless the interstitial atom is smaller than the rest of the atoms in the crystal.

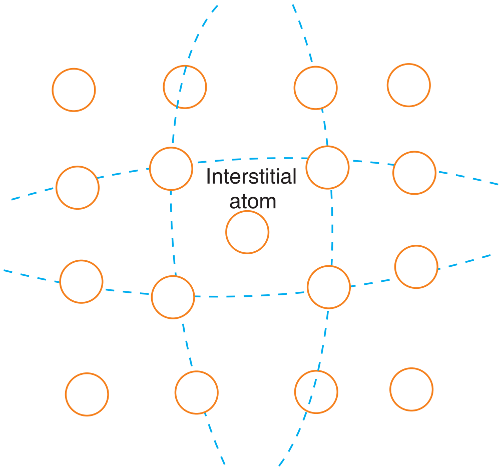


Fig. 77.26

Frenkel Defect. *Frenkel defect* is closely related to interstitialcies. An ion displaced from the lattice into an interstitial site is called a Frenkel defect (Fig. 77.27). When the defect migrates through the crystal, it does so by a sequence of jumps of the kind shown in Fig. 77.27, rather than by the migration of a single ion from one interstitial position to another.

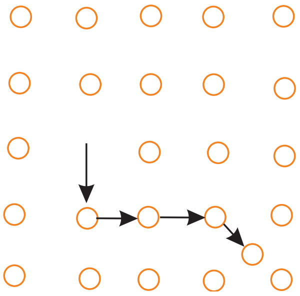


Fig. 77.27

Frenkel defect in Ionic Crystals

In the case of ionic crystals, an ion displaced from the lattice into an interstitial site is called a *Frenkel defect*.

As cations are generally the smaller ions, it is possible for them to get displaced into the void space present in the lattice. Anions do not get displaced like this, as the void space is just too small for their size. A Frenkel imperfection does not change the overall electrical neutrality of the crystal.

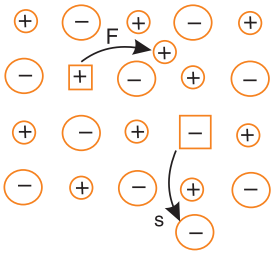


Fig. 77.28

Theoretical and experimental evidence shows that Frenkel defects are most commonly found in silver halides. The density of crystal does not change due to the presence of Frenkel defects because, to the first order of approximation, there is no net change of volume of the crystal.

The Schottky and Frenkel defects in an ionic crystal are shown in Fig. 77.28.

The arrows show how the ions are displaced. In Schottky defect *S*, the ion is removed to the surface of the crystal. In a Frenkel defect *F*, the ion moves to such an interstitial position which is not a normal position of any ion.

The Frenkel and Schottky defects together are called '*Intrinsic defects*'.

77.13.3 Impurities

Impurities give rise to compositional defects. Impurities may be small particles (such as slag inclusions in metals) embedded in the structure, or foreign (metal) atoms in the lattice. Foreign atoms generally have atomic radii and electronic structures differing from those of the host atoms and therefore act as centres of distortion. Impurities may considerably distort the lattice.

There are two types of impurity defects.

(*i*) Substitutional impurity. A substitutional impurity refers to a foreign atom that substitutes for or replaces a parent atom in the lattice (Fig. 77.29).

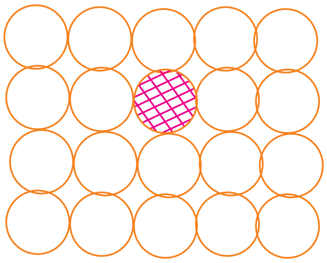


Fig. 77.29

Examples. 1. In ionic solids (e.g., in NaCl), the substitution of Na^+ by Li^+ produces a substitutional impurity.

2. In semiconductor technology, aluminium and phosphorus doped in silicon are substitutional impurities in the crystal.

A controlled addition of impurity to a very pure crystal is the basis of producing many electronic devices.

3. During the production of brass alloy, the zinc atoms are doped in copper lattice. Here the zinc atoms are called substitutional impurities.

(*ii*) Interstitial impurity

An interstitial impurity is a small sized atom occupying the void space in the parent crystal, without dislodging any of the parent atoms from their sites (Fig. 77.30).

An atom can enter the interstitial or void space only when it is substantially smaller than the parent atom.

Example. In FCC iron, the atomic radius of iron atom is 0.225 nm. Carbon atoms with atomic radius 0.0777 nm can occupy the octahedral void spaces in FCC lattice as interstitial impurities.

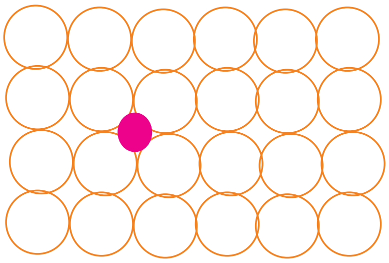


Fig. 77.30

77.13.4 Electronic Defects

Electronic defects are the result of errors in charge distribution in solids.

These defects are free to move in the crystal under the influence of an electrical field. This accounts for some electronic conductivity of certain solids and their increased reactivity.

Example. In zinc oxide (ZnO), the zinc ions occupy interstitials. This leads to a large number of positive charges at interstitials. Suppose in some places zinc ions are missing. Then, at these places, there is a gain of negative charges due to loss of positive charges. Thus a vacancy or an interstitial impurity may produce an excess or deficiency of positive or negative charges.

77.14 To Derive an Expression for the Number of Schottky Defects at a Given Temperature

Fig. 77.31 shows two kinds of Schottky defects most commonly found in crystals, namely, the individual positive and negative vacancy, presumed to form near each other, and a coupled pair.

Consider a perfect crystal composed of equal numbers of positively and negatively charged ions. In order for a cation vacancy to occur, a positive ion must somehow migrate out of its proper position in the structure to the crystal's exterior. If only positive ions migrate out of the crystal and collect on its surface, the surface will become positively charged. This positive surface charge opposes the migration of additional positive ions out of the crystal's interior. Simultaneously, the excess negative charge created inside the crystal is conducive to the formation of negative vacancies. In the absence of external forces, therefore, the number of oppositely charged vacancies inside a crystal tends to be equal.

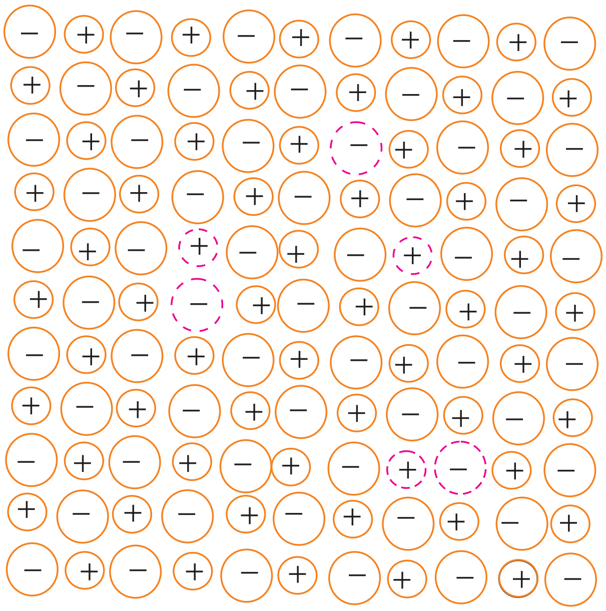


Fig. 77.31

Suppose that the crystal contains a total of N atoms.

Suppose that n Schottky defects are produced by removing n cations and n anions from the crystal's interior.

The different ways in which each kind of ion can be removed is given by

$$\frac{N(N-1)(N-2)...(N-n+1)}{n!} = \frac{N!}{(N-n)!n!}. \quad ...(1)$$

The number of cation and anion vacancies are euqal.

So different ways in which n Schottky defects can be formed is obtained by squaring the expression in Eq. (1).

$$\therefore \quad P = \left[\frac{N!}{(N-n)!n!}\right]^2 \quad ... (2)$$

The creation of n Schottky defects increases the crystal's entropy, according to the Boltzmann relation, by an amount

$$S = k \ln \left[\frac{N!}{(N-n)!n!}\right]^2 \quad ...(3)$$

Let E_p be the energy required to remove a pair of atoms from the crystal's interior to the surface.

Total change in internal energy = $E = nE_p$.

Here, n is the number of vacancy pairs.

This, in turn, produces a change in Helmholtz free energy

$$\begin{aligned} F &= E - TS \\ &= nE_p - kT \ln \left[\frac{N!}{(N-n)!n!}\right]^2 \quad ...(4) \end{aligned}$$

The logarithmic term in Eq. (4) containing factorials can be simplified by using Stirling's formula, $\ln x! = x \ln x - x$.

$$\begin{aligned} \ln \left[\frac{N!}{(N-n)!n!}\right]^2 &\simeq 2[\ln N! - \ln (N-n)! - \ln n!] \\ &= 2[N \ln N - N - (N-n) \ln (\mathrm{N} - n) + (N-n) - n \ln n + n] \\ &= 2[N \ln N - (N-n) \ln (N-n) - n \ln n]. \quad ...(5) \end{aligned}$$

When equilibrium is attained at a given temperature T, the Helmholtz free energy is constant and its first derivative is, therefore, equal to zero.

$$\begin{aligned} \left(\frac{\partial F}{\partial n}\right)_T &= 0 = E_p - 2kT [\ln (N-n) - \ln n] \\ &= E_p - 2kT \ln \frac{N-n}{n} \quad ...(6) \end{aligned}$$

Here the partial differentiation is with respect to n since the total number of atomic positions in the crystal N is not altered.

$$E_p = 2kT \ln \frac{N-n}{n} \quad ...(7)$$

$$\text{or} \quad \frac{N-n}{n} = e^{E_p/2kT} \quad ...(8)$$

The number of Schottky defects in a crystal is much smaller than the number of atoms; that is, $n \ll N$ and $N - n \simeq N$.

Eq. (8) reduces to

$$n \simeq Ne^{-Ep/2kT} \quad ...(9)$$

Eq. (9) is used to determine the approximate number of defects present at any temperature.

77.15 To Derive an Experssion for the Number of Frenkel Defects at a Given Temperature

In Frenkel defect, an atom is transferred from a lattice site to an *interstitial position*, a position not normally occupied by an atom (Fig. 77.32).

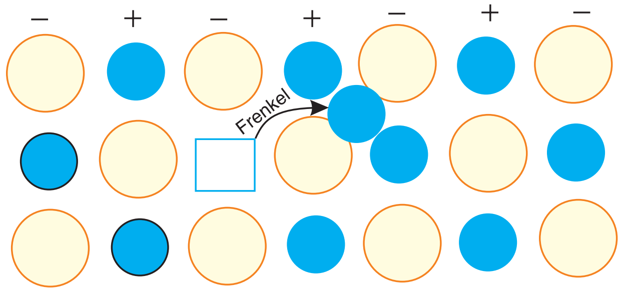


Fig. 77.32

In a perfect crystal, the energy required to displace an atom from its proper position to an interstitial position is E_i.

Suppose there are N atoms in the crystal and N_i interstitial positions in its structure.

The total number of ways in which n Frenkel defcets can be formed is given by

$$P = \frac{N!}{(N-n)!n!}\,\frac{N_i!}{(N_i-n)!n!} \qquad \text{...(1)}$$

The corresponding increase in entropy due to the creation of Frenkel defects is given by

$$S = k\ \ln\left\{\left[\frac{N!}{(N-n)!n!}\right]\times\left[\frac{N_i!}{(N_i-n)!n!}\right]\right\}$$

The amount of energy required to produce n Frenkel defects is

$$E = nE_i.$$

The change in the Helmholtz free energy produced by the creation of n Frenkel defects is

$$F = E - TS$$

$$\therefore \qquad F = nE_i - kT\ln\frac{N!}{(N-n)!n!}\,\frac{N_i!}{(N_i-n)!n!} \qquad \text{...(2)}$$

The factorial terms are simplified by using Stirling's formula

$$\ln x! = x\ln x - x.$$

$$\ln\frac{N!}{(N-n)!n!} + \ln\frac{N_i!}{(N_i-n)!n!} \simeq N\ \ln N + N_i \ln N_i$$

$$-(N-n)\ln(N-n) - (N_i - n)\ln(N_i - n) - 2n\ln n. \qquad \text{...(3)}$$

Substituting (3) into (2), we get

$$F = nE_i - kT\left[N\ln N + N_i\ln N_i - (N-n)\ln(N-n) - (N_i-n)\ln(N_i-n) - 2n\ \ln n\right] \qquad \text{...(4)}$$

Differentiating Eq. (3) with respect to n,

$$\left(\frac{\partial F}{\partial n}\right)_T = E_i - kT\ln\frac{(N-n)(N_i-n)}{n^2} \qquad \text{...(5)}$$

At equilibrium $\qquad (\partial F/\partial n)_T = 0.$

$$\therefore \qquad E_i = kT \ \ln\left[\frac{(N-n)(N_i-n)}{n^2}\right]$$

Now, $N \gg$ and $Ni \gg n$.

$$\therefore \qquad \frac{E_i}{kT} \simeq \ln\frac{NN_i}{n^2}$$

$$= \ln (NN_i) - 2 \ln n$$

Rearranging the terms, we have

$$\ln n = \frac{1}{2}\ln(NN_i) - \frac{E_i}{2kT}$$

$$\therefore \qquad n = (NN_i)^{\frac{1}{2}} e^{-E_i/2kT} \qquad \text{... (6)}$$

In silver halides, the most prevalent defects below 700 K are Frenkel defects.

77.16 To Derive an Expression for the Number of Vacancies at a Given Temperature

In all crystals vacancies are present. The main cause for these defects is thermal agitation. Suppose E_v is the energy required to take an atom from a lattice site inside the crystal to a lattice site on the surface. Suppose there are N atoms.

The amount of energy required to produce n number of isolated vacancies is

$$E = nE_v \qquad \text{...(1)}$$

The total number of ways in which n number of atoms can be removed out of N number of atoms in a crystal on to the surface is

$$P = \frac{N!}{(N-n)!n!} \qquad \text{...(2)}$$

The increase in entropy due to formation of n vacancies is

$$S = k \ \ln \ P = k \ \ln\left\{\frac{N!}{(N-n)!n!}\right\}$$

The change in the Helmholtz free energy produced by the creation of n vacancies is

$$F = E - TS$$

$$\therefore \qquad F = nE_v - kT \ \ln\left\{\frac{N!}{(N-n)!n!}\right\}$$

$$F = nE_v - kT\,[\ln N! - \ln (N-n)! - \ln n!]$$

The factorial terms are simplified by using Stirling approximation,

$$\ln x! = x \ln x - x.$$

$$F = nE_v - kT\,[N \ln N - N - (N-n) \ln (N-n) + (N-n) - n \ln n + n]$$

$$F = \text{nEv} - \text{kT}\,[N \ln N - (N-n) \ln (N-n) - n \ln n] \qquad \text{...(3)}$$

At thermal equilibrium, free energy is constant and minimum.

Differentiating Eq. (3) with respect to n,

$$\left(\frac{\partial F}{\partial n}\right)_T = 0 = E_v - kT\left[0-(N-n)\frac{-(1)}{(N-n)} - (-1)\ln(N-n) - n\frac{1}{n} - 1\ln n\right]$$

$$E_v = kT\,[1 + \ln (N-n) - + \ln n]$$

$$E_v = kT \ln\left[\frac{(N-n)}{n}\right]$$

$$\frac{N-n}{n} = \exp\left(\frac{E_v}{kT}\right)$$

$$n \equiv (N-n)\exp\left[\frac{-E_v}{kT}\right]$$

If $n << N$, n can be neglected.

$$\therefore \qquad n \equiv N \exp\ N \exp\left[\frac{-E_v}{kT}\right]$$

If $E_v = 1\text{eV}$ and $T = 1000$ K, then

$$\frac{n}{N} = e^{-11.6} = 10^{-5}.$$

The equilibrium concentration of vacancies decreases as the temperature decreases.

77.17 Line Defects

Line defects are called dislocations. These are one-dimensional imperfections in the geometrical sense. A *dislocation* may be defined as a disturbed region between two substantially perfect parts of a crystal (Fig. 77.33).

Dislocation is a line defect in a crystal structure whereby a part-plane of atoms is displaced from its symmetrically stable positions in the array. It is surrounded within the structure by an extensive elastic strain field and its associated stresses.

There are two basic types of dislocations :

1. Edge dislocation, and
2. Screw dislocation

77.17.1 Edge Dislocation

Fig. 77.34 (a) shows a perfect crystal.

The top sketch shows a three dimensional view of a perfect crystal.

The bottom sketch shows the arrangement of atoms on its front face.

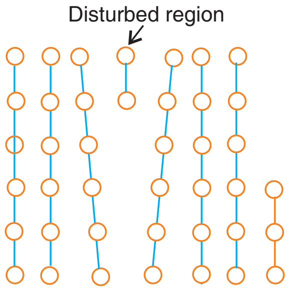
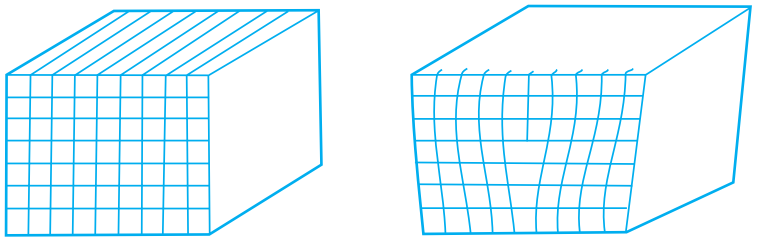


Fig. 77.33



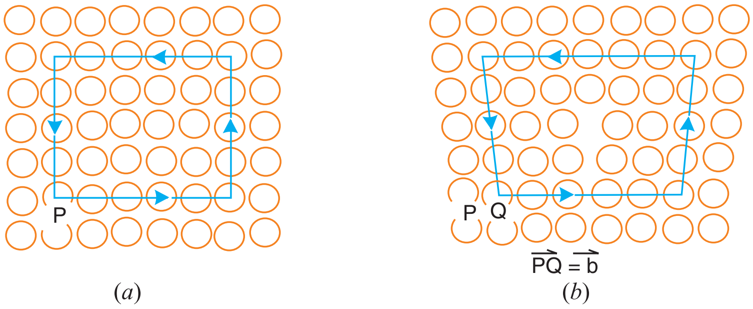


Fig. 77.34

Fig. 77.34 (b) shows an imperfect crystal.

The top sketch shows a three dimensional view of an imperfect crystal with an edge dislocation.

The bottom sketch shows the arrangement of atoms on its front face. The atoms above the edge of the extra plane are squeezed together and are in a state of compression. The bond lengths have been compressed to smaller than the equilibrium value. Just below the edge of the extraplane, the atoms are pulled apart and are in a state of tension. Here the bond lengths have been stretched to above the normal values. This distorted configuration extends all along the edge into the crystal. There is an extra energy due to the distortion in the region immediately surrounding the edge of the incomplete plane. As the region of maximum distortion is centred around the edge of the incomplete plane, this distortion represents a line imperfection and is called an *edge dislocation*.

The vector $\vec{b} = \overrightarrow{PQ}$ connecting the end point Q with the starting point P is the Burgers vector of the dislocation.

Edge dislocation can be classified as *positive edge dislocation* and *negative edge dislocation*.

An edge dislocation involves an extra row of atoms, either above (positive sign) or below (negative sign) the slip plane (Fig. 77.35).

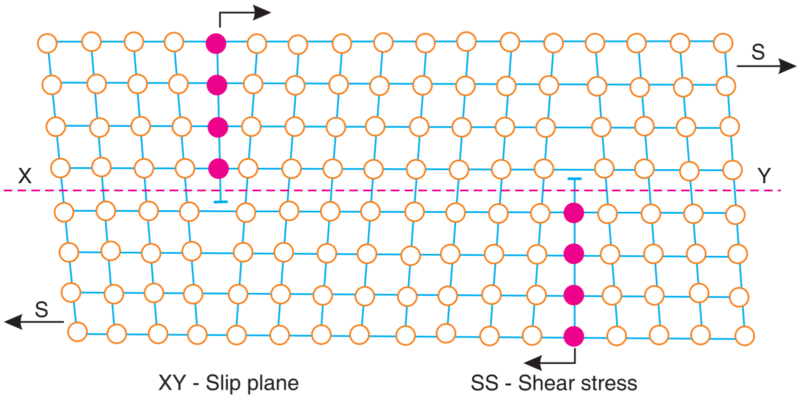


Fig. 77.35

The horizontal line XY is called the slip plane.

(*i*) Positive dislocation

If an extra plane of atoms is above the line XY, the edge dislocation is said to be positive. It is denoted by the symbol $\perp$.

(*ii*) Negative dislocation

If an extra plane of atoms is below the line XY, the dislocation is said to be negative. It is denoted by the symbol T.

77.17.2 Movement of □ and T Dislocations Under Shear Stress □

The presence of an extra plane of atoms means that adjacent atoms are displaced elastically. Consequently, from both sides elastic forces are exerted on the dislocation. These forces balance out, so that it is easy to move the dislocation from one position to another.

Under a shear stress sense a positive dislocation (⊥) moves to the right and a negative dislocation (*T*) to the left (Fig. 77.35).

Slip caused by the movement of edge dislocation

Fig. 77.36 shows how edge dislocation (*D*) gets glided along the slip plane under the action of shear stress (*S*) and how the slip is caused by the movement of an edge dislocation.

As the dislocation glides out of the crystal completely, it produces a *slip step* of one atom width at the edge of the crystal.

The edge dislocation is particularly useful in explaining slip in plastic flow during mechanical working.

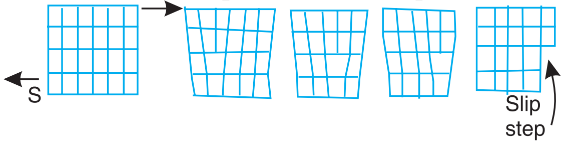


Fig. 77.36

77.17.3 Burger's Vector

The Burger's vector indicates how much and in what direction the lattice above the slip plane appears to have been shifted with respect to the lattice below the slip plane. The Burger's vector is perpendicular to the edge dislocation.

Burger's Vector marks the magnitude and direction of the strain component of dislocation.

Method of determining Burger's Vector for Edge Dislocation

Fig. 77.37 shows the method of determining the Burger's Vector.

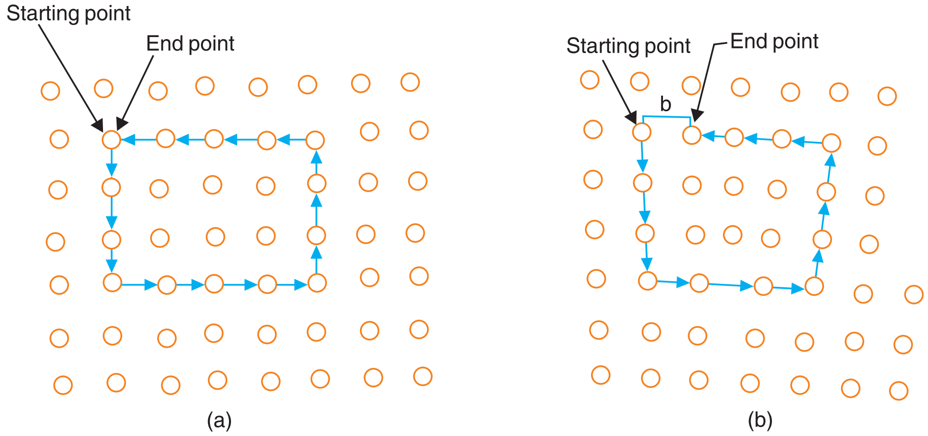


Fig. 77.37

The Burger's circuit is formed by proceeding through the undisturbed region surrounding a dislocation in steps which are integral multiples of a lattice translation. The loop is completed by going an equal number of translations in a positive sense and negative sense in a plane normal to the dislocation line.

(*i*) The Burger's loop closes upon itself if it does not enclose a dislocation [Fig. 77.37 (a)].

(*ii*) The Burger's circuit surrounding an edge dislocation is shown in Fig. 77.37 (b). The end point does not coincide with the starting point.

Therefore, the vector $\vec{b}$ connecting the end point with the starting point is the Burger's Vector of the dislocation .

77.17.4 Screw Dislocation

Screw dislocation results from a displacement of the atoms in one part of a crystal relative to the rest of the crystal, forming a spiral ramp around the dislocation line.

Fig. 77.38 shows what happens when one part of the crystal is displaced relative to the rest of the crystal and the displacement terminates within the crystal.

The row of atoms marking the termination of the displacement is the screw dislocation. *E F* indicates the dislocation line.

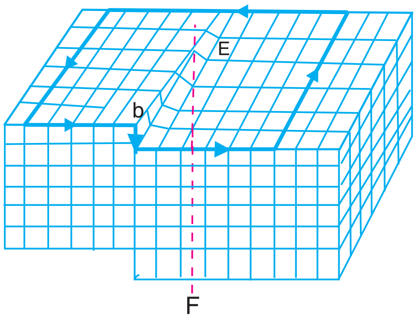


Fig. 77.38

In screw dislocation, the Burger's vector lies *parallel* to the dislocation line along the axis of a line of atoms in the same plane. The Burger's vector (Fig. 77.39) determines the magnitude and direction of the screw dislocation. The screw dislocation may be thought of as produced by cutting the crystal partway through with a knife and shearing it parallel to the edge of the cut by one atom spacing. A screw dislocation transforms successive atom planes into the surface of a helix.

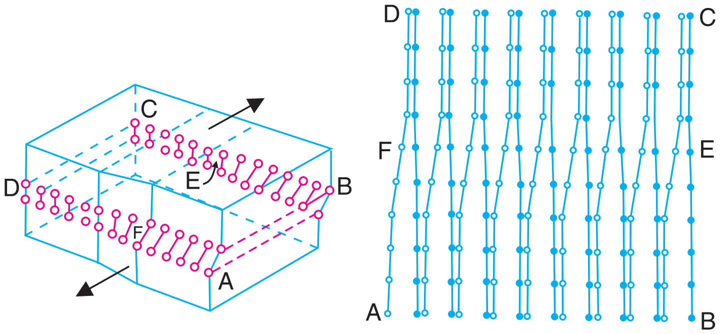


Fig. 77.39

A screw dislocation is sketched in Fig. 77.39.

A part *ABEF* of the slip plane has slipped in the direction parallel to the dislocation line *EF*.

A screw dislocation may be visualized as a helical arrangement of lattice planes, such that we change planes on going completely around the dislocation line.

Fig. 77.40 shows another view of a screw dislocation. The broken vertical line that marks the dislocation is surrounded by strained material.

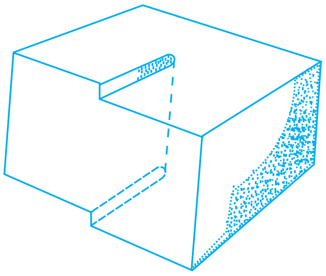
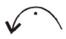
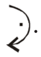


Fig. 77.40

77.17.5 Differences between Edge Dislocation and Screw Dislocation

The following Table gives the differences between edge dislocation and screw dislocation.

Edge Dislocation	*Screw Dislocation*
1. An edge of an atomic plane is formed internal to the crystal.	Only a distortion of the lattice cells in the immediate vicinity is produced.
2. The Burger's vector of an edge dislocation is perpendicular to the dislocation line and lies in the slip plane.	The Burger's vector of a screw dislocation is parallel to the dislocation line.

Edge Dislocation	*Screw Dislocation*
3. An edge dislocation moves (in its slip plane) in the direction of the Burger's vector (slip direction).	A screw dislocation moves (in the slip plane) in a direction perpendicular to the Burger's vector (slip direction).
4. An edge dislocation involves an extra row of atoms either above (positive sign) or below (negative sign) the slip plane. Symbols : ⊥ (or) *T*.	In the screw dislocation, the distortion follows a helical or screw path. The pitch of the screw may be left-handed or right-handed. Symbols :  or .
5. The forces required to form and to move an edge dislocation are smaller in their values.	The forces required to form and move a screw dislocation are somewhat greater than those required to initiate an edge dislocation.
6. Speed of movement of an edge dislocation is greater than that of a screw dislocation.	Speed of movement of a screw dislocation is less than that of an edge dislocation.
7. The edge dislocation is particularly useful in explaining slip in plastic flow during mechanical working.	Screw dislocation is especially useful in explaining crystal growth as well as slip in plastic deformation.

77.18 Surface Defects

In a surface defect, the crystalline irregularity extends in two dimensions. Surface defects are also called plane defects.

Surface imperfections refer to regions of distortions that lie about a surface having a thickness of a few atomic diameters.

Surface imperfections of a structural nature arise from a change in the stacking of atomic planes on or across a boundary. The change may be one of the orientations or of the stacking sequence of the planes. There are two types of surface imperfections.

1. External surface imperfections.
2. Internal surface imperfections.

77.18.1 External Surface Imperfections

The external surface of a crystal is an imperfection in itself, as the atomic bonds do not extend beyond the surface. Although we may visualise an external surface as simply a terminus of the crystal structure, the atoms on the surface cannot be compared with the atoms within a crystal. The surface atoms have neighbours on one side only, while atoms inside the crystal have neighbours on either side of them (Fig. 77.41). Since the external surface atoms are not entirely surrounded by others, they possess higher energy than that of internal atoms. This energy of the surface atom, for most metals, is of the order 1J/m^2.

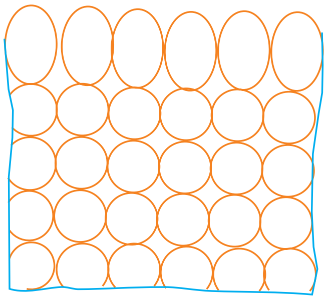


Fig. 77.41

77.18.2 Internal Surface Imperfections

Internal surface imperfections arise from a change in the stacking of atomic planes across a boundary. The change may be one of the orientation or of the stacking sequence of the planes.

The important internal surface imperfections are given below :

(*i*) Grain boundaries

(*ii*) Tilt boundaries and twist boundaries

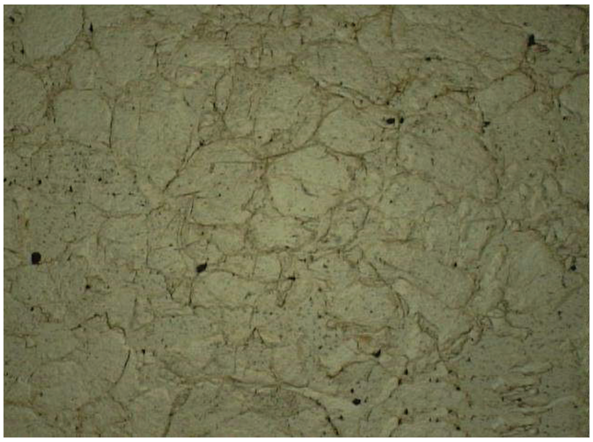
(*iii*) Twin boundaries

(*iv*) Stacking faults and

(*v*) Ferromagnetic domain walls.

77.18.3 Grain Boundaries

Most of the engineering materials are polycrystalline in nature. Grain boundaries are those surface imperfections which separate crystals or grains of different orientation in a polycrystalline aggregation during nucleation or crystallisation. During solidification or during recrystallization of polycrstalline crystals, new crystals are randomly oriented with respect to one another. They grow by the addition of atoms from the adjacent regions and eventually impinge on each other. When two crystals impinge in this manner, the atoms that are caught in between the two crystals are being pulled by each of the two crystals to join its own configuration. They can join neither crystal due to the opposing forces. Therefore, they take up a compromise position. The thickness of this region is only a few atomic diameters, because the opposing forces from neighbouring crystals are felt by the intervening atoms only at such short distances. The boundary region is called a crystal boundary or a grain boundary. The crystal orientation changes sharply at the grain boundary (Fig. 77.42).



Grain Boundaries.

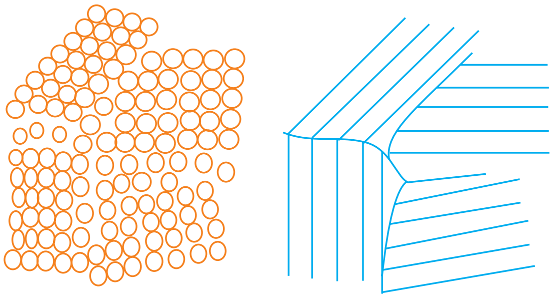


Fig. 77.42

Definition. *Grain boundaries are those planar imperfections in polycrystalline materials that separate crystals (grains) of different orientation.*

A grain boundary is formed when two growing grain surfaces meet.

For example if we take a piece of iron or copper, it is not in a single crystalline form, but it consists of many small interlocking crystals or grains having random orientations. The boundary between these interlocking crystals must have a structure common to both the orientations and structures of the adjacent crystals. Therefore, the grain boundary forms a discontinuity in the

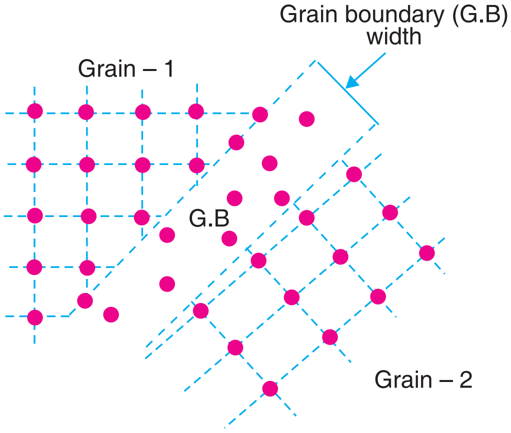


Fig. 77.43

periodicity of the lattice of crystallite or grain. This type of lattice imperfection is called *grain boundary*. The crystal orientation changes sharply at the grain boundary (Fig. 77.42).

The grain boundary has a definite thickness of 2 to 10 or more atomic distances. A grain boundary thickness of a few atoms has been confirmed by the field ion micrography. Boundaries between grains in commercial metals are often wider, because impurities commonly concentrate at the grain boundaries during solidification.

In grain boundaries, the atomic packing is imperfect. At the grain boundary between two adjacent grains, there is a transition zone which is not aligned with either grain (Fig. 77.43).

Most atoms at the boundaries are located in highly strained and distorted positions and their free energy is higher than that of the atoms in the regular, undisturbed part of the crystal lattice.

High angle grain boundary. If the misorientation angle (θ), between the grains is greater than 10°– 15°, the resultant grain boundaries are called *High-angle grain boundaries*. These imperfections in the grain boundary enable the microscopist to see them, for in a transparent crystalline material they may scatter light and in opaque material they can be etched chemically.

Low angle grain boundary. Low angle grain boundaries have angular misorientation of the order of few degrees but less than 10°.

77.18.4 Tilt Boundaries

Tilt boundaries is another surface imperfection. This is called low-angle boundary as the orientation difference between two neighbouring crystals is less than 10°.

Tilt boundary may be regarded as an array of edge dislocations (⊥) (Fig. 77.44).

By rotation of an axis in the boundary it is possible to bring the axis of two bordering grains into coincidence, *i.e.*, a **tilt boundary**.

The angle of tilt is given by

$$\tan\theta = \frac{b}{D} \text{ or } \theta = \frac{b}{D}$$

(because in tilt boundary, the degree of misorientation of the grain boundary or angle of tilt, *i.e.*, θ is less than 10°).

D is the dislocation spacing and *b* is the length of Burger's vector.

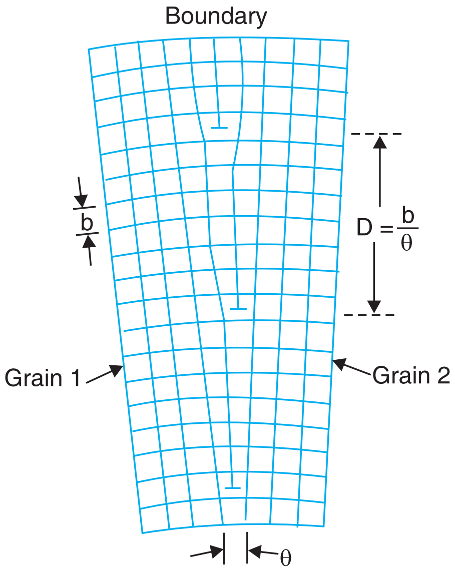


Fig. 77.44

77.18.5 Twist Boundaries

Twist boundaries are the second class of low angle boundaries resulting from the set of screw dislocations. In a twist boundary, the rotation is about an axis normal to the boundary.

77.18.6 Twin Boundaries

Surface defects which separate two orientations that are mirror images of one another are called *twin boundaries*. Twin boundaries occur in pairs, such that the orientation change introduced by one boundary is restored by the other (Fig. 77.45).

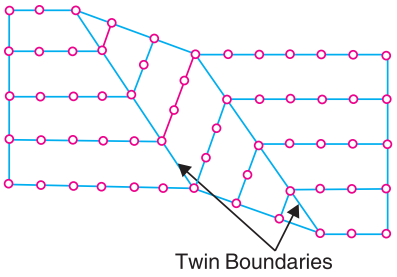


Fig. 77.45

The region between the pair of boundaries is called the twinned region. Twin boundaries are easily identified under an optical microscope. Twins which form during the process of recrystallization are called *annealing twins*. Twins which form during plastic deformation of the material are called *deformation twins*.

77.18.7 Stacking Faults

A stacking fault is a surface imperfection that results from the stacking of one atomic plane out of sequence on another, while the lattice on either side of the fault is perfect. The stacking fault is a discrepancy in the packing sequence of the layers of atoms (*although all the lattice sites are occupied*).

It is usually produced during the growth of the crystals and plastic deformation of crystals.

Example : In the case of close packed FCC structure the stacking sequence can be written as ABC, ABC.....In that sequence it is possible in one atom layer 'A', the atoms are not positioned properly in a small region and hence deviates from the sequence, relative to the atoms of the layers above and below giving a defect, since now there is a sequence of .....BCBC.....which belongs to HCP structure instead of ....ABC ABC..... .

Fig. 77.46 shows the stacking fault in a FCC metal. So one may conclude that stacking fault may arise when there is only small dissimilarity (electrostatically) between the stacking sequences of close-packed planes in FCC and HCP metals.

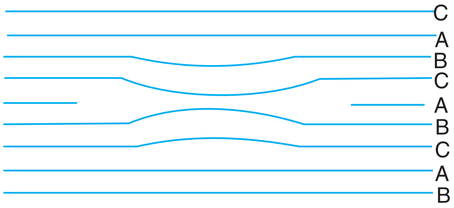


Fig. 77.46

77.18.8 Ferromagnetic Domain Walls

When two ferromagnetic regions differ from one another only in the direction of magnetisation, the boundary between them is an imperfection and is called a ferromagnetic domain wall. These domain walls determine the magnetic properties of ferromagnetic materials.

77.18.9 Volume Defects

Volume defects such as cracks may arise when there is only small electrostatic dissimilarity between the stacking sequence of closed packed planes in metals.

When clusters of atoms are missing, a large vacancy or void is created which is also a volume imperfection.

Foreign particle inclusions, large voids or non -crystalline regions which have the dimensions of at least 1 to 3 nm are also called volume imperfections.

77.19 Effects of Crystal Imperfection

1. The addition of impurities in the case of semiconductors increases the electrical conductivity. But in the case of metals, it decreases the electrical conductivity.

2. Colour Centres

The colour of a crystal is due to the selective absorption of some component of the visible spectrum by certain imperfections, called *colour centres*, which are present in the crystal. These imperfections may be interstitial impurity atoms such as transition-metal ions. Alternatively, they may be vacancies in the structure produced by deviations from stoichiometry so that there are present excess positive ions accompanied by negative vacancies or excess negative ions accompanied by positive vacancies. Thus, when excess Zn is present in ZnO it takes on a yellow colour. Excess lithium in LiF turns the crystal pink. Excess potassium in KCl makes the crystal appear violet.

3. The imperfections account for

(*i*) Flow and fracture characteristics
(*ii*) Crystal growth
(*iii*) Diffusion mechanisms
(*iv*) Creep characteristics of real metals and alloys
(*v*) Annealing and precipitation
(*vi*) Oxidation and corrosion
(*vii*) Yield strength, fracture strength, plasticity, thermal conductivity, dielectric strength, etc.

4. The defects influence the properties of the solids in the following different ways :

(*i*) The defects may scatter conduction electrons in a metal, increasing its electrical resistance by several percent in many pure metals and much more in alloys.
(*ii*) Some defects, even if present sparingly, decrease the strength of the crystal.
(*iii*) Pure salts having impurities and imperfections are often coloured.
(*iv*) Hysteresis loss of ferromagnets.

5. Imperfections do not always affect the properties of materials adversely. There are many situations in which controlled amount of imperfection can bring about specific characteristics desired in a material. For example, carbon steel is a material whose properties are improved by imperfections. The interstitial carbon atoms which may be thought of as imperfections in the cubic crystal structure of iron, increase the mechanical strength of iron.

MAGNETIC MATERIALS

77.20 Ferrites and their Applications

In *ferrimagnetic materials* (also called ferrites) such as $MnFe_2O_4$, the magnetic moments of adjacent ions are antiparallel and of unequal strength (Fig. 77.46A). So there is a finite net magnetisation. Ferrites exhibit spontaneous magnetization below the ferromagnetic Neel temperature T_{FN}.

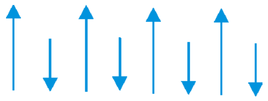


Fig. 77.46A

The general chemical formula of ferrites may be written as XY_2Z_4, where X is a divalent negative ion, Y is Fe^{3+}, and Z is mostly the divalent oxygen ion, O^{2-}. A familiar example is Fe_3O_4 (or $Fe^{2+}Fe^{3+}_2 O^{2-}_4$).

Structure of Ferrites: Ferrites usually crystallize in the spinel structure. The mineral *spinel* is $MgAl_2O_4$. Its unit cell contains 56 ions, 24 of which are iron ions and 32 are oxygen ions. Among the 24 Fe ions, 16 are ferric (Fe^{3+}) ions and 8 are ferrous (Fe^{3+} ions). Out of the 16 ferric ions, 8 are in the tetrahedral sites (where Fe ion is surrounded by 4 oxygen ions) and the remaining 8 in octahedral sites (where it is surrounded by 6 oxygen ions). As the magnetic moments of tetrahedral and octahedral structures are in opposite directions, the magnetic moment due to ferric ions is zero. The net magnetic moment of the system is therefore due to 8 ferrous ions occupying the octahedral sites. By replacing the ferrous ions in Fe_3O_4 with metallic ions like Ni, Mn, Mg, etc., other ferrites may be obtained.

Properties:

1. These are metal oxides, but not metals.
2. These materials exhibit hysteresis property.

3. They are insulators with very high electrical resistivity (~ 10^{12} Ω cm). As a consequence, there will be no eddy current loss as usually noted with iron, at high frequencies.
4. High microwave dielectric constant (~ 10 to 12) (low dielectric loss).
5. A high magnetic permeability (500 – 1000) for the mixed Ni-Zn ferrites ($Ni_{0.36}Zn_{0.64}Fe_2O_4$) at low frequencies and ~ 10 at high frequencies (≥ 300 MHz).

Applications of Ferrites

1. Ferrites are used to produce low frequency ultrasonic waves by magnetostriction principle. Further these are used in the electromechanical transducers.
2. Ferrite rods are used in radio receivers (particularly in medium wave coil) to increase the sensitivity and selectivity of receiver.
3. Ferrites like Nickel Zinc ferrites are used as cores in audio and T.V. transformers.
4. Since for ferrites eddy current loss and hysteresis loss are small at microwave frequencies, these are widely used in non-reciprocal microwave devices like *gyrator, circulator* and *isolator*.
5. Ferrites are also used in digital computers and data processing circuits. Normally here ferrites with rectangular hysteresis loops are used as magnetic storage elements.
6. Based on nonlinear tensor permeability property, ferrites can be used in devices for power limiting and harmonic generation.

77.21 Ferromagnetic Domains

A ferromagnet has a spontaneous magnetic moment — a magnetic moment even in zero applied magnetic field. The existence of a spontaneous moment suggests that electron spins and magnetic moments are arranged in a regular manner (Fig. 77.47).

The magnetization is defined as the magnetic moment per unit volume.

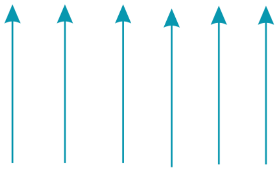


Fig. 77.47

If *domains* (regions magnetized in different directions) are present, the magnetization refers to the value within a domain.

Weiss introduced the concept of *ferromagnetic domains*. He postulates that any specimen consists of a large number of small regions or domains which are spontaneously magnetized. In the demagnetized state, the magnetization vectors of these domains are oriented in such a way that the specimen does not possess a resultant magnetic moment. Fig. 77.48 shows a single crystal domain arrangement for zero resultant magnetic moment. Magnetization occurs only when some of these vectors are caused to align themselves more or less parallel to an applied field .

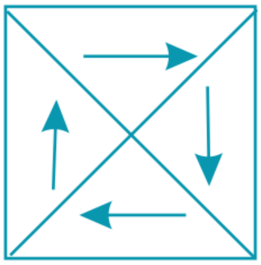


Fig. 77.48

The increase in the gross magnetic moment of a ferromagnetic specimen in an applied magnetic field takes place by two independent processes :

(*i*) By the Motion of Domain Walls

In weak fields, the volume of domains [Fig. 77.49(b)] favourably oriented with respect to the field increases at the expense of unfavourably oriented domains. The magnetization changes by means of domain boundary displacements and so the domains change in size.

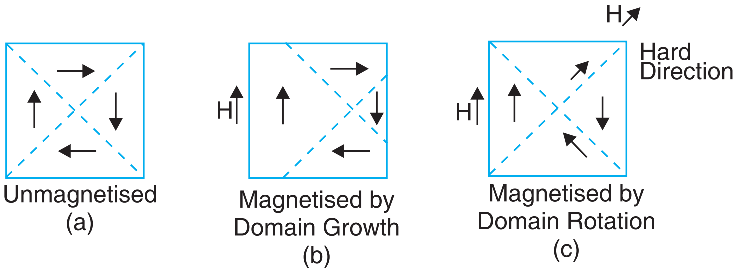


Fig. 77.49

(*ii*) By Rotation of Domains

In strong applied fields, the domain magnetization rotates toward the direction of the field [Fig. 77.49 (c)]. The magnetization changes by means of rotation of direction of magnetization.

Fig. 77.50 depicts the magnetization curve, showing the dominant magnetization processes in the different regions of the curve.

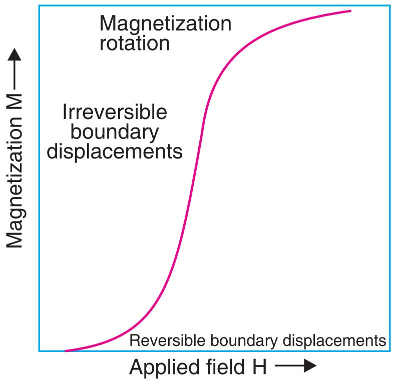


Fig. 77.50

77.22 Four Types of Energy Involved in the Process of Domain Growth

Domain structure occurs in order to minimize the total energy of the ferromagnetic substance. The total energy comprises the sum of the *exchange* energy, the *anisotropy* energy, and the *magnetic* energy.

(*i*) *Exchange Energy*

It is the energy associated with the quantum mechanical coupling that aligns the individual atomic dipoles within a single domain. It depends upon the interatomic distance. It arises from interaction of electron spins. The exchange energy of two electrons may be written in the form

$$U = -2J\,\vec{s}_1 \cdot \vec{s}_2 \qquad \text{... (1)}$$

Here, J is the exchange integral and is related to the overlap of the charge distributions of the atoms *1*, *2*. $\vec{s}_1$, $\vec{s}_2$ are the electron spins of atoms *1*, *2*.

(*ii*) *Anisotropy Energy*

There is an energy in a ferromagnetic crystal which directs the magnetization along certain crystallographic axes called directions of easy magnetization. This energy is called the *magnetocrystalline or anisotropy energy*.

Definition. *The excess of energy required to magnetize a specimen in a particular direction over that required to magnetize it along the easy direction is called crystalline anisotropy energy.*

Iron is a cubic crystal. The cube edges are the directions of easy magnetization. Magnetization curves for single crystal iron are shown in Fig. 77.51. Along the <100> direction, a very small magnitude of the magnetic field is required to produce a large magnetization. It is

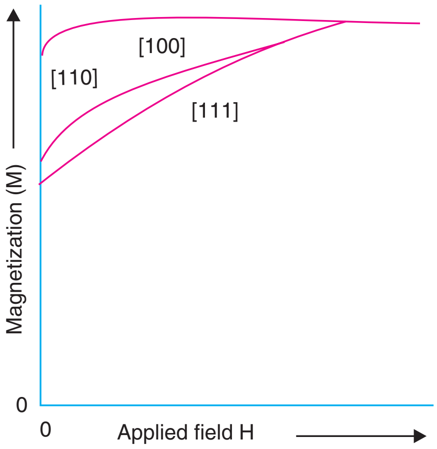


Fig. 77.51

'easy' to magnetize an iron crystal along <100> direction, while it is 'hard' to do so along, say, <111> direction.

To represent the anisotropy energy of iron magnetized in an arbitrary direction with direction cosines α_1, α_2, α_3 referred to the cube edges, we are guided by cubic symmetry. The expression for the anisotropy energy must be an even power of each α_i, provided opposite ends of a crystal axis are equivalent magnetically, and it must be invariant under interchanges of the α_i among themselves.

The lowest order combination satisfying the symmetry requirements is $\alpha_1^2 + \alpha_2^2 + \alpha_3^2$, but this is identically equal to unity and does not describe anisotropy effects. The next combination is of the fourth degree: $\alpha_1^2\,\alpha_2^2 + \alpha_1^2\,\alpha_3^2 + \alpha_3^2\,\alpha_2^2$, and then of the sixth degree : $\alpha_1^2\,\alpha_2^2\alpha_3^2$. Thus

$$U_K = K_1(\alpha_1^2\,\alpha_2^2 + \alpha_2^2\,\alpha_3^2 + \alpha_3^2\,\alpha_1^2) + K_2\alpha_1^2\,\alpha_2^2\,\alpha_3^2.$$

Here, K_1, K_2 are constants.

(iii) Domain Wall Energy (Bloch Wall Energy)

The transition region between the two oppositely magnetized domains is called the Bloch wall.

Fig. 77.52 shows the structure of the Bloch wall separating domains. The arrows represent individual atomic magnetic moments. In iron the thickness of the transition region is about 300 lattice constants.

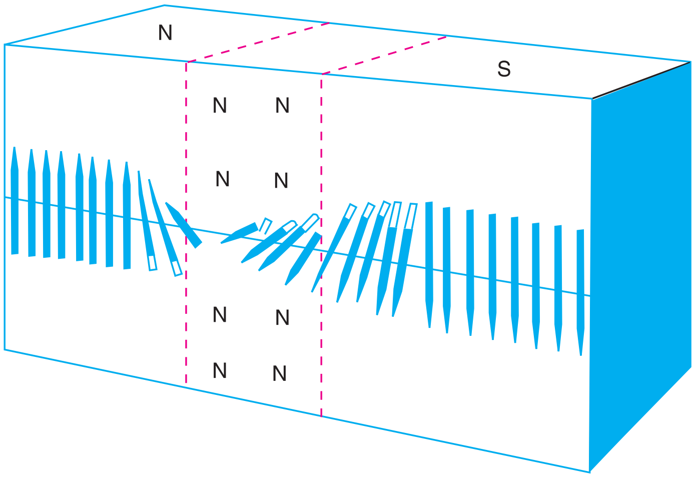


Fig. 77.52

The entire change in spin direction between domains does not occur in one discontinuous jump across a single atomic plane, but takes place in a gradual way over many atomic planes.

Domain wall energy is the sum of contributions from the exchange and crystalline anisotropy energies in the domain wall region.

(iv) Magnetostrictive Energy

Magnetization of a crystal in a certain direction always produces changes in atomic spacing in that direction. This is called *magnetostriction*.

The magnetostrictive energy is the energy due to the mechanical stresses generated by magnetostriction in the domains.

77.23 Reversible and Irreversible Domains

If we apply a small external field on the ferromagnetic specimen, the domain wall is displaced slightly away from the minimum energy. But it returns to the original position when the field is removed. This gives a reversible domain wall movement and the domain is called reversible domain.

In larger external fields, the domain wall may be shifted to a more distant position where the energy curve has passed through a maximum and then diminished. On removing the field, the domain wall cannot cross the energy maximum. So it is unable to return to its initial position. This given an irrversible domain wall movement.

77.24 Explanation of Hysteresis Curve on the Basis of Domain Theory

The Hysteresis of ferromagnetic materials refers to the lag of magnetisation (*M*) behind the magnetising field (*H*).

Technical terms defined by the *B-H* hysteresis loop are shown in Fig. 77.53.

The coercivity H_c is the reverse field that reduces B to zero.

The remanence B_r *is the value of B at H = O.*

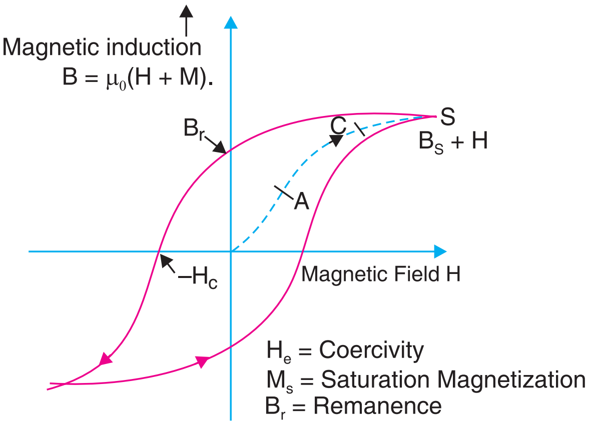


Fig. 77.53

(*i*) When a weak magnetic field is applied, domains where the magnetisation is parallel or at a small angle with the field grow at the expense of those where the magnetisation is antiparallel so that the boundary between domains is displaced. This boundary displacement is reversible and is indicated by the path *OA*.

(*ii*) When the magnetic field becomes stronger, the Bloch wall movement is sharp and is irreversible. The steeper part *AC* of the magnetisation curve is due to larger, irreversible displacements.

(*iii*) Above the knee of the curve (*CS*), magnetization proceeds by rotation of the direction of magnetization of whole domains. Such a process is rather difficult and the increase in magnetisation is relatively slow. At *S*, all the domains are in the field direction and the specimen is said to be saturated.

(*iv*) When the applied field is reduced, there is a little change in the domain structure so that the magnetisation remains quite high, until high reverse fields are applied. Further even when the external field is zero, there is a residual magnetisation in the specimen and that can be destroyed by applying a high reverse field.

Thus the reversible and irreversible domain wall movements give rise to hysteresis in the ferromagnetic materials.

CRYSTAL GROWTH

77.25 Introduction

Nucleation. Crystal growth takes place when a material having preference for a crystalline form is cooled below its melting point slowly. But crystallization will not occur unless there are atoms sufficiently static to act as nucleation centres, to which additional atoms can attach. Crystallites will begin to grow on these nucleation centres. Crystal growth here begins almost simultaneously at many nucleation centres. Each crystallite has a different orientation of its crystalline axes. Crystallites from the many nucleation centres grow until they join, forming grain boundaries and halting growth.

Most synthetic single-crystal technology is based upon growth from a single nucleation region. If even a small single crystal of the material is available, it may be used as a "seed" to which additional atoms can attach. A slow cooling rate is essential for growth of large crystals.

Techniques of Single Crystal Growth

The techniques of crystal growth may be divided into three categories:

(*a*) growth from solutions, (*b*) growth from melts, and (*c*) growth from vapour.

77.26 Crystal Growth from Solution

The growth from solutions takes place in two stages: (*i*) nucleation and (*ii*) growth. Therefore, the first requirement is that of a suitable solvent which can dissolve the solute to an appreciable extent. This may be possible either under normal conditions (atmospheric pressure and room temperature) or high pressures/temperatures. The solvent may be water or melt of a metal.

Nucleation can be carried out from highly supersaturated solutions, so that it can occur spontaneously; but conditions should be such that the solution becomes meta-stable after a few new nuclei have formed. By metastable condition we mean a very small supersaturation, which prevents formation of many nuclei. The growth can then occur over the few 'seed' nuclei already formed. As the growth proceeds, the supersaturation decreases. In order to maintain a reasonable growth rate, the supersaturation has to be kept constant. This can be achieved by (*i*) isothermal evaporation of the solvent, (*ii*) lowering the temperature, (*iii*) adding solute or (*iv*) adding another solvent.

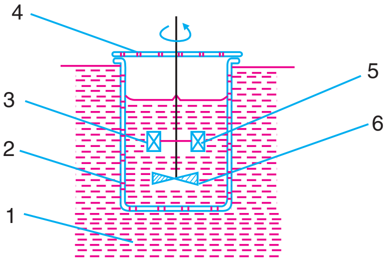
Crystal Growth apparatus

Fig. 77.54 shows the experimental apparatus.

The 'solution' is the liquid in which given substance is dissolved. The dissolving substance - solid or liquid - is known as 'solute' and the bulk liquid is known as 'solvent'.

The seed crystal may be of the same crystalline substance of appropriate size or of a substance which is chemically and crystallographically similar. Any fragment of the solid, no matter how tiny, is a potential seed. It must be a single crystal so that crystal growing from it will also be single.

Many of us are experienced in growing copper sulphate crystals employing this technique.



1. Water bath of the thermostat
2. Crystallisation vessel
3. Solution
4. Perfectly or only partially closing lid
5. Rotably mounted seed crystals
6. Mixer

Fig. 77.54

Advantages

(*i*) Many crystalline compounds produce large single crystals from low-temperature solutions. A great advantage of growing crystals from solutions is that growth is often possible at temperatures far below the melting point of the material. So a lower crystal growth temperature leads to a lower number of defects.

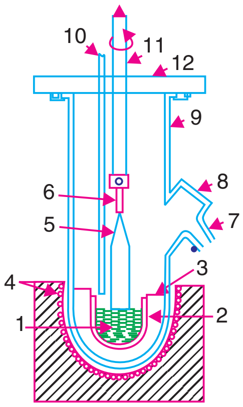
(*ii*) Contamination of the desired material by oxidation is also reduced at the reduced temperatures characteristic of growth from solution.

Disadvantage

Small percentages of impurities in the solvent may produce high concentrations in the crystal. Atoms or molecules of the solvent itself may become entrapped within the rapidly growing crystal.

77.27 Crystal Growth from Melt – Czochralski Method

Fig. 77.55 shows Czochralski's crystal pulling equipment.



1. Si melt
2. Quartz liner
3. Graphite cup
4. Heating elements
5. Growing crystal
6. Seed crystal
7. Hydrogen-argon inlet tube
8. Viewing port
9. Quartz envelope
10. Quartz tube for adding impurities
11. Shaft for pulling and rotating crystal
12. Brass support plate.

Fig. 77.55

Single crystals of semiconductor grade silicon are grown by the Czochralski (CZ) method. A typical CZ apparatus weighs about 20 tonnes, is 7 to 8 metres tall and holds a charge of 100 kg of silicon.

Main Parts

1. A refractory crucible contains the pure material, heated by magnetic induction. The crucible should have a melting point higher than that of silicon.

2. A seed crystal is fastened to a rod, in the desired orientation. The seed is lowered into the surface of the molten material . The temperature of the melt is adjusted to permit atoms to attach to the seed at a rate determined by the temperature and the cooling effect of the rod and the seed. At first the seed is slightly melted to expose pure and relatively defect free material.

3. The rod is rotated about its axis, and at the same time drawn slowly upward at the rate at which crystal growth increases the length of the crystal. If the rod is pulled up slowly, the diameter increases because build-up occurs on the sides as well as on the end of the crystal. Most germanium or silicon crystals are grown to a diameter of about 3 cm. While crystals many centimetres in diameter are possible, larger diameters involve serious problems in temperature uniformity and therefore growth rate, over the growing surface. Thus more defects and stresses are introduced.

4. If desired, the melt in the crucible of the CZ process can be doped with a suitable dopant.

Working

The semiconductor grade polycrystalline silicon is melted in the silica crucible. A seed crystal of the desired orientation is made to touch the surface of the melt. The two commonly used orientations for the axis of the crystal are (111) and (100). The handle holding the seed is slowly pulled upwards at a speed of 50 to 100 mm per h. The temperature T is maintained at $T = T_m$ at the liquid-seed interface, at $T < T_m$ above the interface and at $T > T_m$ below (inside the liquid). This avoids any nucleation of the solid crystal in the liquid and also prevents remelting of the growing crystal. The handle to which the seed is attached is rotated at 6 to 8 rpm and the crucible holding the liquid is rotated in the opposite direction at 10 to 12 rpm. These rotations ensure uniformity in temperature and avoid the appearance of hot or cold spots. By a suitable device, the crucible is gradually lifted upwards to compensate for the falling level of the liquid as the crystal grows. The crystal grows in the same orientation as the seed, without the need for the formation of a grain boundary at the interface.

Advantages

1. The crystal can be observed as it grows and adjustment in temperature and growing rate can be made as needed.

2. The growth rate is quite rapid. Growth rates as high as 10 cm per hour are possible.

77.28 Crystal Growth from Vapour Phase

Single crystals can be grown from the vapour phase by sublimation and chemical vapour deposition (CVD). Crystal growth from vapour phase is suitable for materials which sublime before melting.

Chemical Vapour Deposition (CVD)

Fig. 77.56 shows the apparatus used for the growth of crystals by chemical transport reaction.

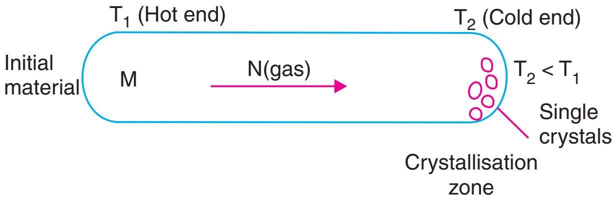


Fig. 77.56

Experimentally chemical transport reactions are carried out in a closed reaction chamber, in which there is a temperature gradient. In an evacuated quartz ampoule, the constituents (*M*) of the required crystal are placed at the hot end and mixed with a low pressure gaseous transport reagent (*N*).

(*i*) *M* and *N* react at the higher temperature to form a volatile species which diffuses along the tube.

(*ii*) At the cold end, the reaction is reversed. The gaseous product decomposes at the lower temperature to deposit material *M*. The transporting agent diffuses to the hotter end and again reacts with the charge.

As a specific example, let us consider the growth of nickel oxide using hydrogen chloride as the transporting agent:

$$\text{At } T_1 \ (\sim 950^0\text{ C})\ \text{NiO(s)} + 2\text{HCl(g)} \rightarrow \text{NiCl}_2\text{(g)} + \text{H}_2\text{O(g)}$$

$$\text{At } T_2 \ (\sim 920°\text{C})\ \text{NiCl}_2\text{(g)} + \text{H}_2\text{O(g)} \rightarrow \text{NiO(s)} + 2\text{HCl(g)}$$

Verneuil Method

The Verneuil technique is an important crystal growth technique from vapour phase. Powders of the crystalline material are thrown into a hot flame or an electric arc. A seed crystal is placed just below the hottest region. Intense heat makes the powder to atomize. As the surface of the growing crystal is brought near the melting point, the atoms or molecules will have high mobility on the growing surface. Synthetic rubies and sapphires are grown by this technique.

MAGNETIC PROPERTIES OF MATERIALS

77.29 Weiss Theory of Paramagnetism

Weiss introduced the concept of *internal molecular field* in order to explain the complicated type of dependence of susceptibility. In a real gas, the molecules are mutually influenced by their magnetic moments and consequently, there should exist within the gas a molecular field. This field produced at any point by all the neighbouring molecules, is proportional to and acting in the same sense as the intensity of magnetization (M). Let this internal molecular field be H_i. Now

$$H_i = \lambda M \qquad \text{...(1)}$$

Here, λ is molecular field coefficient.

Therefore, the net effective field should be

$$H_e = (H + H_i) \qquad \text{...(2)}$$

Here, H is external applied field.

Following the Langevin theory along with this effective field,

$$M = \frac{Nm^2 \mu_0 H_e}{3kT} = \frac{Nm^2 \, \mu_0 \, (H + \lambda M)}{3kT} \quad \text{[Refer Eq. (8) in Langevin theory]}$$

Here, N = number of dipoles/unit volume,

m = magnetic moment of each atomic dipole,

k is Boltzmann's constant and T is the absolute temperature.

Rearranging the terms, we get

$$M\left(1 - \frac{Nm^2 \, \mu_0 \lambda}{3kT}\right) = \frac{Nm^2 \, \mu_0 H}{3kT}$$

Let $C = \mu_0 \, Nm^2/3k$ and $\theta = C\lambda$. Then we get

$$M\left(1 - \frac{\theta}{T}\right) = \frac{CH}{T}$$

$$\therefore \qquad \chi = \frac{M}{H} = \frac{C}{T\left(1 - \frac{\theta}{T}\right)} = \frac{C}{T - \theta} \qquad \text{...(3)}$$

Here C is called the Curie constant.

θ is called *paramagnetic Curie point* or the *Curie temperature.*

Eq. (3) is called *Curie-Weiss law.*

Below Curie temperature ($T < \theta$), susceptibility becomes negative; *i.e.*, paramagnetics would become diamagnetics. However, for most of the paramagnetic substances, Curie temperature is quite low so that a situation for which $T < \theta$ is rare.

77.30 Heisenberg's Interpretation of Internal Field or Weiss Molecular Field

The first theoretical explanation of the large Weiss internal field in ferromagnetic materials was proposed by Heisenberg.

1. The magnetization of ferromagnetic materials arises due to electron spin rather than due to orbital motion.

2. High value of internal fields can be explained in terms of quantum *exchange interactions* between the electrons.

3. An exchange force is caused by a static coulomb interaction and its magnitude is much larger than that of the magnetic interaction.

4. This exchange force appears in the form of spin-spin interaction.

5. The strength of interaction depends upon the interatomic separation and it may change its sign as the separation is varied.

6. Assume in each atom unpaired electrons are in parallel orientation. If the interatomic separation decreases more and more, exchange forces decrease until finally they pass through zero and an antiparallel spin alignment is favoured.

Theory

Consider two ferromagnetic atoms *A* and *B* in the lattice (Fig. 77.57).

Ψ_A and Ψ_B denote the atomic wavefunctions.

1 and 2 are the two unpaired electrons having negative charges.

The interaction potential 'V_{AB}' between the ions of *A* and *B* and electrons 1 and 2 is

$$V_{AB} = \frac{e^2}{4\pi\varepsilon_0}\left[\frac{1}{r_{AB}} + \frac{1}{r_{12}} - \frac{1}{r_{B1}} - \frac{1}{r_{B2}} - \frac{1}{r_{A1}} - \frac{1}{r_{A2}}\right]$$

Here, r_{AB} = distance between atoms *A* and *B*,

r_{B1} = distance between atom *B* and electron 1

r_{B2} = distance between atom *B* and electron 2

r_{A1} = distance between atom *A* and electron 1

r_{A2} = distance between atom *A* and electron 2.

The charge of each ion (atom) is equal to *e*.

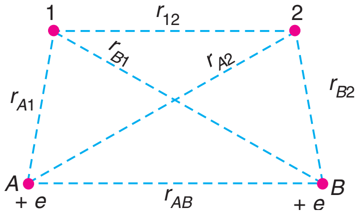


Fig. 77.57

The total exchange interaction energy between spins of unpaired electrons of the system is given by

$$E_{ex} = -2\sum_{ij}(J_{ij})_e\, \vec{S_i}.\vec{S_j}$$

Here $S_i = S_j = 1/2$ = spin magnetic moment.

For parallel spins, $\theta = 0$ and for antiparallel spins $\theta = \pi$.

Here exchange integral $J_e = \int \Psi_A^*(1)\,\Psi_B^*(2)\,V_{AB}\,\Psi_A(1)\,\Psi_B(2)\,d\mathbf{r}_1 d\mathbf{r}_2$

which depends on overlap of atomic orbitals.

The total energy of the system is

$$E_{ex} = K \pm J_e \qquad \text{...(1)}$$

Here, *K* is the coulomb interaction energy and J_e is the exchange integral.

The positive sign in Eq. (1) stands for the state when the spins of the electrons are antiparallel, *i.e.*, non-magnetic state. The negative sign is for the state when the spins of the two electrons are parallel, *i.e.*, for magnetic state.

Fig. 77.58 shows variation of exchange integral with distance ratio r_{AB}/r_0.

r_0 is the orbital radius of unpaired electrons.

A magnetic state is favoured when r_{AB}/r_0 is slightly larger than 3. For Fe, Co and Ni this ratio has the values 3.26, 3.64 and 3.94 respectively.

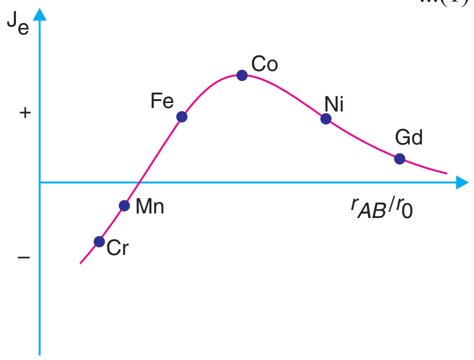


Fig. 77.58

So Fe, Co, and Ni are ferromagnetic. For Cr and Mn this ratio has the value 2.60 and 2.24, *i.e.*, these are not ferromagnetic.

77.31 Quantum Theory of Ferromagnetism

Ferromagnetic substances are those substances which possess a spontaneous magnetic moment *i.e.*, a magnetic moment even in the absence of an applied magnetic field. The Curie temperature θ is the temperature at and above which the spontaneous magnetisation vanishes.

Weiss gave the theory of ferromagnetism on the basis of two hypotheses —

(1) A ferromagnetic specimen contains in general a large number of small regions called domains which are spontaneously magnetised. The spontaneous magnetisation of the specimen is determined by the vector sum of the magnetic moments of the individual domains.

(2) Within each domain the spontaneous magnetisation is due to existence of molecular field which tends to produce a parallel alignment of the atomic dipoles.

Existence of molecular moment leads to co-operation or interaction between the atomic dipoles. Tendency of co-operation is to produce parallel alignment.

To explain this fact Weiss assumed that magnetic field actually acting on the specimen is

$$H_{eff} = H + \lambda M \qquad \text{...(1)}$$

On these assumptions we shall develop the quantum theory of ferromagnetism.

According to quantum theory, since the magnetic moments are quantized, the magnetic dipole moment μ and its component μ_z in the direction of the applied field can not have arbitrary values. We have, in general, a direct relationship between the magnetic dipole moment μ of an atom or ion in free space and its angular momentum J as

$$\mu = -g\,\mu_B\,J \qquad \text{...(2)}$$

μ_B is the Bohr magneton. It is defined as $e \quad /2m$.

The g factor is given by the Lande equation

$$g = 1 + \frac{J(J+1) + S(S+1) - L(L+1)}{2J(J+1)} \qquad \text{...(3)}$$

Here, S and L represent the spin and orbital quantum numbers of the dipole respectively.

The orientations of the magnetic moment μ with respect to the direction of the applied magnetic field are specified by the rule that the possible components of μ along the field direction are given by

$$\mu_z = -g\,\mu_B\,m_J \qquad \text{...(4)}$$

$$m_J = J, (J-1), \ldots -(J-1), -J.$$

For each value of J, m_J can have $(2J+1)$ values which means that the magnetic moments of the atom can have $(2J+1)$ different orientations relative to the field.

The potential energy of the magnetic dipole in the presence of a magnetic field H is

$$U = -\vec{\mu} \,.\, \vec{H} = m_J\,\mu_B\,g\,H_{eff} \qquad \text{...(5)}$$

The number of atoms with a given value of m_J is proportional to $\exp(m_J\,g\,\mu_B\,H_{eff}/kT)$.

Consider a ferromagnetic solid containing N atoms per unit volume each having a total angular momentum number J. Then the magnetization in the direction of the field is given by

$$M = N.\frac{\sum\limits_{m_J=-J}^{J} m_J\,g\mu_B.\exp.(m_J\,g\mu_B H_{eff}/kT)}{\sum\limits_{m_J=-J}^{J} \exp(m_J\,g\mu_B H_{eff}/kT)} \qquad \text{...(6)}$$

Substituting $\frac{g\,\mu_B\,H_{eff}}{kT} = x$ in the above equation, we get

$$M = Ng\,\mu_B \frac{\sum\limits_{m_J=-J}^{J} m_J \exp(m_J\,x)}{\sum\limits_{m_J=-J}^{J} \exp(m_J\,x)} \qquad ...(7)$$

After some algebraic manipulations, Eq. (7) reduces to,

$$M = Ng\mu_B\left[\frac{2J+1}{2}\coth\frac{2J+1}{2}x - \frac{1}{2}\coth\frac{x}{2}\right]$$

Substituting $x = a/J$ in the above equation, we get

$$M = Ng\mu_B\left[\frac{2J+1}{2}\coth\frac{2J+1}{2J}a - \frac{1}{2}\coth\frac{a}{2J}\right]$$

$$= Ng\mu_B J\left[\frac{2J+1}{2J}\coth\frac{2J+1}{2J}a - \frac{1}{2J}\coth\frac{a}{2J}\right]$$

or $$M = Ng\,J\,\mu_B B_J(a) \qquad ...(8)$$

Here $B_J(a)$ is the *Brillouin function* and it is defined as

$$B_J(a) = \frac{2J+1}{2J}\coth\left(\frac{2J+1}{2J}\right)a - \frac{1}{2J}\coth\left(\frac{a}{2J}\right) \qquad ...(9)$$

From the two substitutions made in the above treatment, we have

$$\frac{a}{J} = \frac{g\mu_B\,H_{eff}}{kT}$$

or $$a = \frac{gJ\mu_B\,H_{eff}}{kT} = \frac{g\,J\mu_B}{kT}(H + \lambda M) \qquad ...(10)$$

For spontaneous magnetisation $H = 0$. Eq. (10) becomes

$$a = \frac{gJ\mu_B\,\lambda M}{kT} \qquad ...(11)$$

$\therefore$ $$M(T) = \frac{akT}{gJ\mu_B\lambda} \qquad ...(12)$$

As $T \to 0$ or $a \to \infty$, $B_J(a) \to 1$; the magnetic moments align themselves parallel to the field and the magnetization M becomes the *saturation magnetization,* $M_S(0)$. Thus, from Eq. (8), we get

$$M_S(0) = Ng\,J\,\mu_B \qquad ...(13)$$

From Eqs. (12) and (13), we obtain,

$$\frac{M(T)}{M_S(0)} = \frac{akT}{\lambda N g^2 J^2 \mu_B^2} \qquad ...(14)$$

Eqs. (8) and (13) give

$$\frac{M(T)}{M_S(0)} = B_J(a) \qquad ...(15)$$

$\frac{M(T)}{M_S(0)}$ must simultaneously satisfy both equations (14) and (15). So its magnitude at a given temperature is obtained graphically as the intersection of the two $\frac{M(T)}{M_S(0)}$ versus a plots (Fig. 77.59).

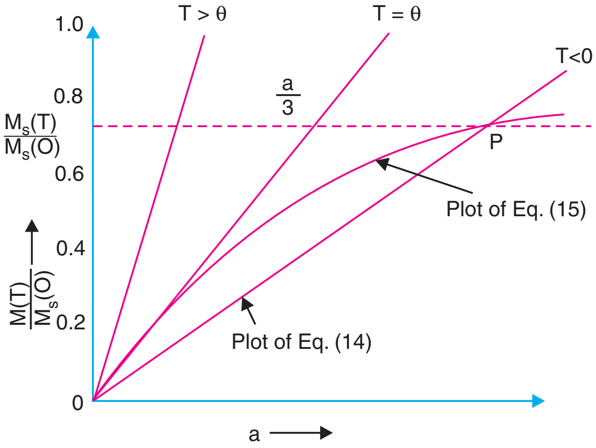


Fig. 77.59

(*i*) For $T < \theta$ (curie temperature), spontaneous magnetization results.

(*ii*) For $T \geq \theta$, the two curves do not intersect and there is no spontaneous magnetization.

DIELECTRIC MATERIALS

77.32 Fundamental definitions in dielectrics

1. Dielectric constant (ε_r)

Dielectric constant is the ratio between the absolute permittivity of the medium (ε) and the permittivity of free space (ε_0)

$$\text{Dielectric constant } (\varepsilon_r) = \frac{\text{Absolute permittivity of the medium } (\varepsilon)}{\text{Permittivity of free space } (\varepsilon_0)}$$

2. Electric polarisation

The process of producing electric dipoles by an electric field is called polarisation in dielectrics.

When an electric field is applied to a crystal or a glass containing positive and negative charges, the positive charges are displaced in the direction of the field, while the negative charges are displaced in the opposite direction. This displacement produces local dipoles throughout the solid.

3. Polarisation vector *P*

*The polarisation vector **P** is the dipole moment per unit volume of the dielectric material.*

$$P = N\bar{\mu}$$

Here, N = number of molecules per unit volume and

$\bar{\mu}$ = average dipole moment per molecule.

4. Electric displacement vector *D*

The electric displacement vector ***D*** is given by

$$\boldsymbol{D} = \varepsilon \boldsymbol{E} = \varepsilon_0 \varepsilon_r \boldsymbol{E} \qquad \text{...(1)}$$

Here, $\boldsymbol{E}$ = electric field strength,

ε_0 = permittivity of free space (*i.e.,* vacuum),

ε_r = relative permittivity of the material.

Relation between $\boldsymbol{D}$ and $\boldsymbol{P}$ is

$$\boldsymbol{D} = \varepsilon_0 \boldsymbol{E} + \boldsymbol{P} \qquad \text{...(2)}$$

$$\varepsilon_0 \varepsilon_r \boldsymbol{E} = \varepsilon_0 \boldsymbol{E} + \boldsymbol{P}$$

$$\therefore \quad \boldsymbol{P} = \varepsilon_0 (\varepsilon_r - 1)\boldsymbol{E}$$

5. Electric susceptibility (χ_e)

When a dielectric material is placed in an electric field, it becomes electrically polarised. For isotropic dielectrics, the polarisation vector $\boldsymbol{P}$ is proportional to the electric field $\boldsymbol{E}$. Thus,

$$\boldsymbol{P} = \varepsilon_0 \chi_e \boldsymbol{E}$$

The constant χ_e is called the 'electric susceptibility' of the dielectric material.

$$\chi_e = \frac{P}{\varepsilon_0 E} = \frac{\varepsilon_0(\varepsilon_r - 1)\, E}{\varepsilon_0 E}$$

$$\therefore \quad \chi_e = (\varepsilon_r - 1)$$

77.33 Different Types of Electric Polarisation

1. Electronic polarisation

The displacement of the positively charged nucleus and the (negative) electrons of an atom in opposite directions on the application of an electric field results in *electronic polarisation*.

Fig. 77.60 shows the electronic polarisation occuring in an atom due to the applied electric field. On applying a field, the electron cloud around the nucleus shifts towards the positive end of the field. As the nucleus and the centre of the electron cloud are separated by a certain distance, dipole moment is created within each atom.

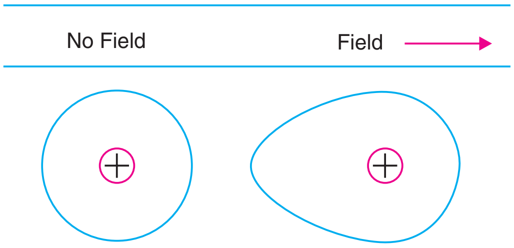


Fig. 77.60

The induced dipole moment (μ) is proportional to the field strength (E).

$$\mu \propto E$$

or

$$\mu = \alpha_e E$$

The constant of proportionality α_e is the electronic polarisability of the atom. The electronic polarisability α_e is proportional to the volume of the atom. It is independent of temperature.

Calculation of electronic polarizability (α_e).

(*i*) No electric field

Consider an atom of a dielectric material of atomic number Z [Fig. 77.61 (*a*)]. Ze is the nuclear charge. The electrons of charge ($-Ze$) are distributed uniformly throughout the sphere of radius R. When there is no external electric field, the centres of the electron cloud and nucleus of the atom are same and one.

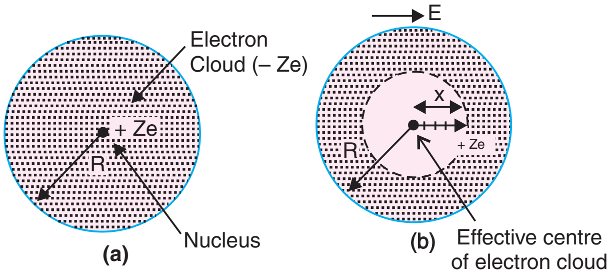


Fig. 77.61

(*ii*) With electric field

When a d.c. electric field E is applied, the nucleus moves towards the field direction and the electron cloud moves in the opposite direction of the field. [Fig. 77.61 (*b*)]. The electron cloud centre gets a displacement x with respect to the centre of the nucleus. The displacement $x << R$. We assume that the spherical shape of the electron cloud is not altered on the application of the electric field.

When the nucleus and electron cloud are shifted from their equilibrium position, an attractive coulomb force is created and it will tend to maintain the original equilibrium position. But the Lorentz force will tend to separate the nucleus and electron cloud of that atom from their equilibrium position. Finally a new equilibrium will be reached when these two forces are equal and opposite.

At equilibrium, *Lorentz force* = *Coulomb force*.

The Lorentz force between nucleus and the electron cloud = $-Ze \times E = -ZeE$ ...(1)

The charge density of the charged sphere = $\dfrac{-Ze}{\frac{4}{3}\pi R^3}$

$$\left.\begin{array}{l}\text{Total number of negative charges}\\ \text{in the sphere of radius } x\end{array}\right\} = \frac{-Ze}{\frac{4}{3}\pi R^3} \times \frac{4}{3}\pi x^3 = -Ze\frac{x^3}{R^3} \qquad ...(2)$$

$$\left.\begin{array}{l}\text{The coulomb force between the}\\ \text{nucleus and the electron cloud}\end{array}\right\} = \frac{\frac{-Ze\,x^3}{R^3}(Ze)}{4\pi\,\varepsilon_0 x^2} = \frac{-Z^2e^2x}{4\pi\,\varepsilon_0 R^3}$$

Under equilibrium conditions, the coulomb attractive force and the Lorentz repulsive force are equal and opposite. Thus

$$ZeE = \frac{Z^2e^2x}{4\pi\,\varepsilon_0 R^3}$$

or

$$E = \frac{Zex}{4\pi\,\varepsilon_0 R^3} \qquad ...(3)$$

Thus, the displacement of the electron cloud (x) is proportional to the applied electric field E.

Induced dipole moment $= \mu = Zex$ ...(4)

By definition $\mu = \alpha_e E$ ...(5)

From Eqs. (4) and (5), $E = \dfrac{Zex}{\alpha_e}$ ...(6)

From Eqs. (3) and (6), we get

$$\frac{Zex}{\alpha_e} = \frac{Zex}{4\pi\,\varepsilon_0 R^3}$$

$$\therefore \qquad \alpha_e = 4\pi\,\varepsilon_0\,R^3 \qquad ...(7)$$

2. Ionic polarisation

The ionic polarisation is due to the displacement of cations and anions in opposite directions and occurs in an ionic solid (Fig. 77.62). The displacement of ions is independent of temperature.

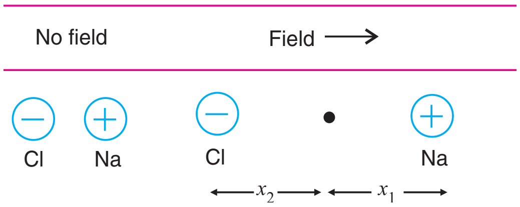


Fig. 77.62

Suppose an electric field is applied in the $+ve\ x$ direction.

Then the + ve ions move to the right by $x_1$and – ve ions move to the left by x_2.

m is the mass of the + ve ion.

M is the mass of the – ve ion.

We assume that there are one cation and one anion in each unit cell of the ionic crystal.

Due to ionic displacement the resultant dipole moment per unit cell,

$$\mu = e(x_1 + x_2) \qquad \text{...(1)}$$

Here, x_1 = shift of + ve ion and x_2 = shift of – ve ion with respect to their equilibrium positions.

When the electric field is applied, the restoring force (F) produced is proportional to the displacement (x).

Restoring force on + ve ion, $F = \beta_1 x_1$ ...(2)

Restoring force on – ve ion, $F = \beta_2 x_2$ ...(3)

Here, β_1 and β_2 are restoring force constants.They depend upon the masses of ions and angular frequency of the molecule in which ions are present. Therefore, under equilibrium,

$$F = \beta_1 x_1 = \beta_2 x_2$$

$$x_1 = \frac{F}{\beta_1} = \frac{eE}{m\omega_0^2} \qquad \begin{bmatrix} \because F = eE \\ \beta_1 = m\omega_0^2 \end{bmatrix}$$

Here, ω_0 = natural frequency of the ionic molecule.

Similarly for – ve ion,

$$x_2 = \frac{F}{\beta_2} = \frac{eE}{M\omega_0^2}$$

$$\therefore \qquad x_1 + x_2 = \frac{eE}{\omega_0^2}\left(\frac{1}{m} + \frac{1}{M}\right) \qquad \text{...(4)}$$

$$\mu = e(x_1 + x_2) = \frac{e^2 E}{\omega_0^2}\left(\frac{1}{m} + \frac{1}{M}\right)$$

$$\therefore \qquad \alpha_i = \frac{\mu}{E} = \frac{e^2}{\omega_0^2}\left(\frac{1}{m} + \frac{1}{M}\right) \qquad \text{...(5)}$$

Thus the ionic polarisability 'α_i' is inversely proportional to the square of the natural freqency of the ionic molecule and to its reduced mass. It is independent of temperature.

3. Orientational Polarisation

In methane molecule (CH_4), the centre of the negative and positive charges coincide. It has no permanent dipole moment.

However in CH_3Cl, the positive and negative charges do not coincide. Here chlorine attracts the bonding electrons to itself more strongly than hydrogen. Therefore, even in the absence of an electric field, this molecule has dipole moment. When an electric field is applied, these molecules tend to align themselves in the applied field as shown in Fig. 77.63.

Hence the electric dipole moment increases. The polarisation due to such alignment is called *orientational polarisation*.

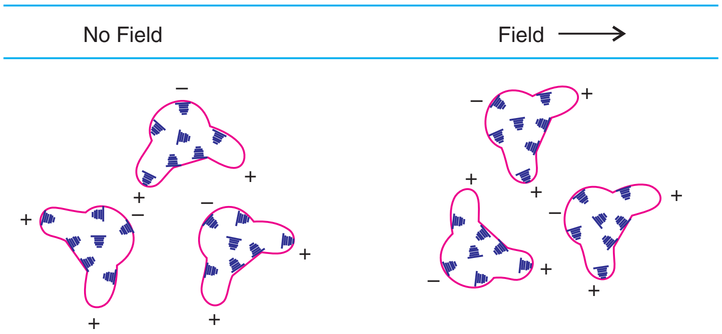


Fig. 77.63

Orientational polarisation can be shown to be

$$P_0 = N\bar{\mu} = \text{N}\,\mu^2 E/3kT = \text{N}\alpha_0 E.$$

$\therefore$ orientational polarisability, $\alpha_0 = \dfrac{P_0}{NE} = \dfrac{\mu^2}{3kT}.$

Thus the orientational polarisability α_0 is inversely proportional to absolute temperature of the material.

4. Space -charge polarisation

Space-charge polarisation occurs due to the accumulation of charges at the electrodes or at the interfaces in multiphase dielectrics (Fig. 77.64). It is found in ferrites and semiconductors. The ions diffuse over appreciable distances in response to the applied field, giving rise to a redistribution of charges in the dielectric medium.

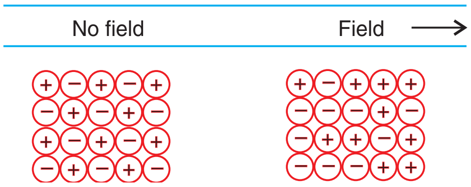


Fig. 77.64

77.34 Total Polarisation

Since the value of space-charge polarisability is very small when compared with other types of polarisabilities, it can be neglected.

The total polarisability of a gas 'α' can be written as

$$\alpha = \alpha_e + \alpha_i + \alpha_0$$

$$= 4\pi\,\varepsilon_0\,R^3 + \frac{e^2}{\omega_0^2}\left(\frac{1}{M} + \frac{1}{m}\right) + \frac{\mu^2}{3kT}$$

The total polarisation P can be written as

$$P = N\,\alpha\,E = NE\left\{4\pi\,\varepsilon_0\,R^3 + \frac{e^2}{\omega_0^2}\left(\frac{1}{M} + \frac{1}{m}\right) + \frac{\mu^2}{3kT}\right\}$$

This equation is known as *Langevin - Debye equation*.

77.35 Effect of Temperature on Polarisation

1. The electronic and the ionic polarisabilities are independent of temperature.

2. The effect of increase of temperature brings about a higher degree of randomness in the molecular orientation in the material. This affects the tendency of permanent dipoles to align along the field direction. Hence, *orientation polarisation decreases with the increase in temperature*.

3. In space charge polarisation, increase in temperature facilitates the diffusion of ions. Thermal energy may also aid in overcoming the activation barrier for the orientation of polar molecules in the direction of the field. Hence, *space charge polarisation increases with increase of temperature*.

77.36 Dielectric Loss

When a dielectric is subjected to the a.c. voltage, the electrical energy is absorbed by the material and is dissipated in the form of heat. This dissipation of electrical energy in the dielectric is called dielectric loss.

Since this involves heat generation and heat dissipation, this assumes a dominating role in high voltage applications.

Expression for dielectric loss (or loss tangent)

(*i*) When a.c. voltage is applied to perfect insulator like vacuum or purified gas, there is no consumption of energy and the charging current leads the applied voltage by 90° [Fig. 77.65 (*b*)].

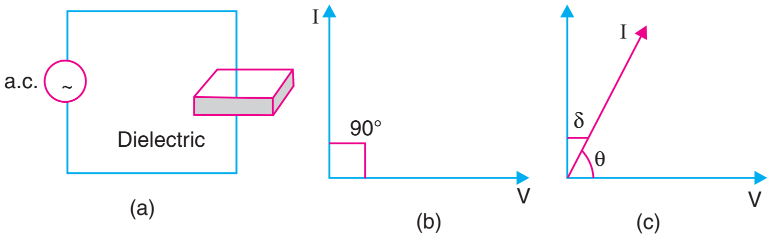


Fig. 77.65

Power loss, $P = VI \cos\theta = VI \cos 90° = 0$.

Thus the ideal dielectrics have no dielectric loss.

(*ii*) The phase angle between charging current and applied voltage is less than 90° in commercial dielectrics. Complementary angle $\delta = 90 - \theta$ is called the *dielectric loss angle [Fig. 77.65(c)]*. The loss angle is a measure of the power dissipated in each cycle.

For a dielectric having a capacitance of C and having a voltage V applied to it at a frequency f Hz, the dielectric power loss is given by

$$P = VI \cos\theta$$

Since $I = \dfrac{V}{X_c}$ where $X_c = \dfrac{1}{2\pi f C}$ = capacitive reactance to A.C.

and $\cos\theta = \cos(90 - \delta) = \sin\delta$,

$$P = \frac{V^2}{X_c} \sin\delta = 2\pi f C V^2 \sin\delta$$

Since δ is very small, $\sin\delta = \tan\delta$.

$\therefore$ Dielectric power loss, $P = V^2\, 2\pi f C \tan\delta$

Here $\tan\delta$ is called *power factor* of the dielectric. The power loss is dependent on $\tan\delta$ as long as the other factors like voltage, frequency and capacitance are constants. Thus the power loss is found to depend on the frequency and value of applied voltage.

77.37 Determination of Dielectric Constant of a Dielectric Material

The Schering bridge is used for measuring the dielectric constant of dielectric materials.

Principle. It is based on measurement of the capacitance of the capacitor without the dielectric material (C_1) and the capacitance of the capacitor with the dielectric material (C'_1).

The connections are shown in Fig. 77.66.

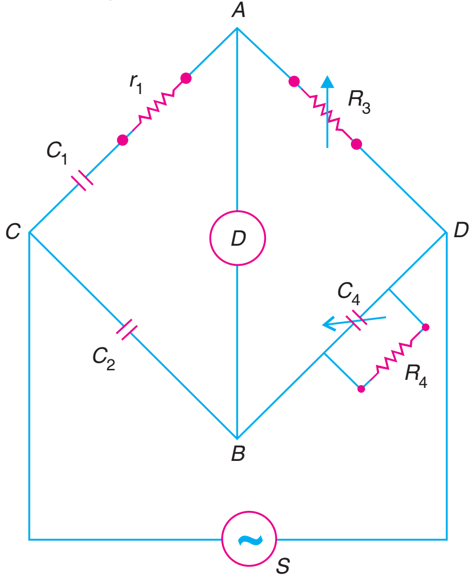


Fig. 77.66

C_1 = Capacitor whose capacitance is to be measured
r_1 = A series resistance representing the loss in the capacitor C_1
C_2 = Standard capacitor
R_3 = Non-inductive variable resistance
R_4 = Non-inductive resistance;
C_4 = Variable capacitor
D = A.C. null detector;
S = High frequency oscillator.

(*i*) First without inserting dielectric inside C_1, the bridge is balanced by varying C_4 and R_3. When current flowing through the detector becomes zero,

$$\frac{\left(r_1 + \dfrac{1}{j\omega C_1}\right)}{R_3} = \frac{\dfrac{1}{j\omega C_2}}{\left(\dfrac{R_4}{(1 + j\omega R_4 C_4)}\right)}$$

$$\left(r_1 + \frac{1}{j\omega C_1}\right)R_4 = \frac{R_3(1 + j\omega R_4 C_4)}{j\omega C_2}$$

$$r_1 R_4 - \frac{jR_4}{\omega C_1} = \frac{-j R_3}{\omega C_2} + \frac{R_3 R_4 C_4}{C_2}$$

Equating real and imaginary parts, we get

$$r_1 = \frac{R_3 C_4}{C_2}$$

$$C_1 = \frac{C_2 R_4}{R_3}$$

Since R_4 and C_2 are fixed, the dial of R_3 is calibrated to read the capacitance C_1 directly.

(*ii*) The dielectric specimen should be in the size of C_1 (in its area) and is inserted between the plates of C_1. Now once again the bridge is balanced. Now the dial reading in R_3 will give the value of new capacitance C'_1.

Then $\frac{C_1'}{C_1} = \varepsilon_r$ = dielectric constant of the specimen.

We can determine the dielectric constant at different frequencies by changing the frequency of the oscillator.

77.38 Dielectric Breakdown

When a dielectric loses its insulation resistance and permits large currents to pass through it, it is said to breakdown.

Whenever the electric field strength applied to a dielectric exceeds a critical value, very large current flows through it. The dielectric loses its insulating property and becomes conducting. This phenomenon is called dielectric breakdown.

There are five important types of dielectric breakdown.

1. Intrinsic breakdown

In a dielectric, the charge displacement increases with increasing field strength. Beyond a critical value of the field strength, there is an electric breakdown due to the physical deterioration of the dielectric material. The *dielectric strength* is defined as the breakdown voltage per unit thickness of the material.

$$\text{Dielectric strength} = \frac{\text{Dielectric breakdown voltage}}{\text{Thickness of dielectric}}.$$

When the applied electric field is large, some of the electrons in the valence band cross over to the conduction band across the large forbidden energy gap. They become conduction electrons and give rise to large conduction currents. The liberation or movement of electrons from valence band is called internal field emission of electrons. This breakdown is called the *Intrinsic breakdown* or *Zener breakdown*.

Avalanche breakdown

Since the electric field is over the order of 10^6 V/cm, the electrons are accelerated to very high velocity, colliding with the other atoms and molecules thereby releasing more electrons and holes breaking covalent bonds between them. The number of bonds broken and hence number of charge carriers released increases enormously with time and finally dielectric breakdown occurs. This type of breakdown is called *Avalanche breakdown*.

This requires relatively large electric fields. Hence this breakdown occurs in thin samples.

Characteristics:

1. This kind of breakdown requires relatively large electric fields.
2. Mostly this kind of breakdown occurs in thin samples.

2. Thermal breakdown

Thermal breakdown occurs in a dielectric when the rate of heat generation is greater than the rate of dissipation. Energy due to the dielectric loss appears as heat. If the rate of generation of heat is larger than the heat dissipated to the surrounding, the temperature of the dielectric increases which eventually results in local melting. Once melting starts, that particular region becomes highly conductive, enormous current flows through the material and dielectric breakdown occurs.

Characteristics:

1. It occurs only at high temperatures.

2. The strength of the electrical field to create dielectric breakdown depends upon the size and shape of the insulation sample.

3. The breakdown time is of the order of few milliseconds.

4. Since the dielectric loss is directly proportional to frequency, the electric field strength to create dielectric breakdown will be smaller for alternating fields and higher for D.C. fields.

3. Electrochemical breakdown

Electrochemical breakdown has a close relationship with thermal breakdown. When temperature rises, mobility of ions increases. Hence electrochemical reaction takes place. When ionic mobility increases, leakage current also increases. This results in dielectric breakdown. The chemical reaction reduces the insulation resistance and finally creates the dielectric breakdown. This type of breakdown is called *electrochemical breakdown*.

Characteristics

1. Electrochemical breakdown is determined by the leakage current, density of ions, temperature and permanent dipoles in the material.

2. Electrochemical reactions are accelerated by high temperatures. So to avoid breakdown, insulating materials should not be operated at high temperatures.

4. Discharge breakdown

Discharge breakdown is classified as internal or external.

(*i*) *Internal breakdown* occurs when the insulator contains occluded gas bubbles. Since gases require smaller ionisation potential than solids, the gaseous atoms ionise first and the gaseous ions bombard the solid dielectric causing electrical deterioration. This can occur at low voltages if there are large number of occluded gas bubbles in the insulating material.

(*ii*) *External breakdown* is caused by a glow or corona discharge. Such discharges are normally observed at sharp edges of electrodes. Carbon is formed on the electrodes due to these discharges. So the damaged areas become conducting leading to power arc and complete failure of the dielectric. Dust and/or moisture on the surface of the dielectric may also cause external discharge breakdown.

When discharge takes place at a point, the surrounding places are burnt and hence their electrical properties are affected. Thus the life of the insulation of an insulating material depends upon the number of discharges which are taking place inside the material. That is, it depends upon the frequency of the applied voltage.

5. Defect Breakdown

The surface of the dielectric materials may have defects such as cracks, porosity and blow holes etc. Then impurities like dust or moisture may collect at these discontinuities (defects). This leads to breakdown called *defect breakdown.*

77.39 Properties and Applications of Different Types of Insulating Materials

Insulating materials are classified into three categories.

1. Solid insulating materials : Mica, Porcelain, asbestos, bakelite, rubber, PVC, polythene, glass, cotton, silk, wood and ceramics..

2. Liquid insulating materials; Varnish, transformer oil, cable oil, and silicon liquids.

3. Gaseous insulating materials: Air**,** nitrogen, inert gases and sulphur hexafluoride.

1. Solid insulating materials

(*i*) Mica

It is a mineral compound of silicate of aluminium with silicates of soda potash and magnesia.

Properties:

1. It is crystalline in nature.
2. It can be easily split into very thin flat sheets.
3. It is rigid, tough and strong. It is not affected by moisture .
4. Its dielectric constant varies between 5 and 7.5.
5. Its dielectric strength varies between 700 and 1000 kV / mm .
6. It has high dielectric strength and low power loss.

Uses:

1. Mica sheets are used as insulator between commutator segments.
2. Thick mica tape is used for taping the stator coils of high voltage alternators.
3. It is used in electrical irons, hot plates and toasters for insulation purposes.
4. It is used as a dielectric material for high frequency applications.

(*ii*) Ceramics

They are generally non-metallic inorganic compounds such as silicates, aluminates, oxides, carbides, borides, nitrides and hydroxides.

Properties:

1. Ceramics are hard, strong, dense and brittle.
2. They are completely stable at high temperatures. They are not affected by moisture and by chemical action except with strong acids and alkalies.
3. They may be either crystalline or amorphous.
4. The ceramics have excellent dielectric and mechanical properties. The dielectric constant of the commonly used ceramics varies between 4 and 10.

Uses:

1. The ceramics are widely used as insulators for switches, plug holders, cathode heaters, vacuum type ceramic metal seals etc.
2. They are very much used as dielectric materials in capacitors. These capacitors withstand very high temperature and very high voltage.
3. **Refractory materials** are the materials which do not deterioate or disintegrate, when they are subjected to high temperatures. Alumina is an important refractory material. It is primarily made of aluminium oxide. It has high compressive strength, absorbs less moisture and is excellent in heat conduction. It is used in high temperature appliances like furnaces.

(*iii*) PVC materials

When acetylene and hydrogen chloride are combined in presence of catalyst at temperature of about 50°C, polyvinyl chloride (PVC) resin is produced.

Properties:

1. They have good mechanical and electrical properties.
2. They are non - corrosive.

Uses:

1. They are widely used in insulation for wires and cables.

2. PVC films, tapes, and sheets are commonly used as insulation for dry batteries, conductors and cables.

(*iv*) Rubber

They are organic polymers and may be natural or synthetic. Natural rubber has limited applications because of its poor stability at wide temperature range. The synthetic rubbers are produced artificially by copolymerisation of isobutylene and isoprene.

Properties:

1.The rubbers have good electrical and thermal properties.

2. Rubber possesses good dielectric properties and high tensile strength.

The dielectric constant of rubber varies between 2.5 and 5.

Uses:

1. Rubber is widely used as an insulating material for electric wires, cables, tapes, coatings, transformers, motor winding etc.

2. It is used in the construction of storage battery housings and panel boards.

(*v*) Glass

Glass is an inorganic material made by the fusion of different oxides like silicon oxide, zinc oxide and magnesium oxide. Silica glass, Pyrex and fibre glass are the important glass insulating materials.

Properties:

1. It is brittle and hard. It is insoluble in water.

2. It has low dielectric loss and has good mechanical strength.

3. It is highly chemical resistant to most corrosive agents. It is a good insulator with good appearance.

Uses:

1. Glass is used as dielectric in capacitors.

2. It is used as an insulator in radio and television tubes, electrical lamps and laminated boards.

3. Toughened glass is used for insulation in extra high voltage lines having voltages above 100 kilovolts.

(*vi*) Asbestos

It is a naturally occurring mineral material of fibrous structure. Asbestos generally consists of magnesium silicate composition.

Properties:

1. It has high dielectric loss and low dielectric strength.

2. It can withstand very high temperature (400°C).

Uses:

1. Asbestos finds extensive use in electrical machines because of its ability to withstand very high temperature.

2. It is used for making of cloth tapes and paper boards for insulation purposes.

3. It is used for covering/insulating the wire in electric heating devices, ovens, electric irons, etc.

2. Liquid insulating materials

Liquid insulating materials can be divided into three groups.

(i) Mineral insulating oils : Transformer oil, cable oil, capacitor oil, etc.

(ii) Synthetic insulating oils : Askarels, aroclors, sovol and sovtol.

(iii) Miscellaneous insulating oils : Vegetable oils, vaseline and silicon liquids.

(*i*) Mineral insulating oils

These oils are obtained from crude petroleum by distillation and have high oxidation resistance and good thermal stability.

Properties and uses

Transformer oil is the important mineral insulating oil with high dielectric strength, viscosity and cooling properties. This oil is used for insulation and cooling of transformer. It transfers heat from windings and core to the cooling surfaces by convection . It maintains the insulation of the windings.

(*ii*) Synthetic insulating oils

Now-a-days synthetic oil is used as an insulator in transformers in the place of transformer oil (mineral oil) because synthetic oils are very much resistant to oxidation and to fire hazards. Due to longer life and safety in operating conditions, synthetic oil is used as coolant and insulator in H. V. transformers. Particularly askarels have thermal stability upto 110°C and are manufactured from chlorinated hydrocarbons.

(*iii*) Miscellaneous insulating oils

Silicon liquids have stability upto 200°C and are costly. The dielectric strength of these liquids is same as that of mineral oils and power factor is very low. They are used in H.V. Transformers. Further these are used to increase the surface resistivity of ceramic insulators. Vaseline has high viscosity and high dielectric constant. It is used for impregnation of papers used in capacitors.

3. Gaseous insulating materials

(*i*) Air

It is the most important insulating material available in nature. The dielectric constant of air increases linearly with the increase of pressure. It is used in air capacitors. The power loss is practically zero. It is used for insulation between the two conductors of high voltage transmission lines. Air can be used as insulation only in the low voltage applications since at higher field strengths air may get ionised.

(*ii*) Nitrogen

Nitrogen is chemically inert. It prevents oxidation and reduces the rate of deterioration. In oil filled transformers, the nitrogen is used to replace oxidizing atmosphere. This is also used in capacitors and in cables under pressure.

(*iii*) Inert gases

They are used in electronic tubes and discharge tubes as insulators.

(*iv*) Sulphur hexafluoride

It is produced when sulphur is burnt in fluorine atmosphere. It has high dielectric strength. It has superior cooling properties than those of air and nitrogen. It has high chemical stability upto 100°C and is used in transformers and electric switches. Further it is also used in Vande Graff generator, Voltage stabilizer and X-ray apparatus for insulation purposes.

PART – XI

ASTROPHYSICS

78. Astrophysics
79. Nanotechnology

CHAPTER

78

ASTROPHYSICS

AT A GLANCE

78.1 Introduction
78.2 Classification of Stars — The Harvard Classification System
78.3 Hertzsprung — Russel Diagram
78.4 Luminosity of a Star
78.5 Stellar Evolution
78.6 White Dwarfs
78.7 Electrons in a White Dwarf Star
78.8 Chandrasekhar Limit
78.9 Neutron Stars
78.10 Black Holes
78.11 Supernova Explosion
78.12 Photon Diffusion Time
78.13 Gravitational Potential Energy of a Star
78.14 Internal Temperature of a Star
78.15 Internal Pressure of a Star

78.1 Introduction

Astrophysics is a branch of astronomy that treats of the physical properties of celestial bodies, such as luminosity, size, mass, density, temperature, and chemical composition, and their origin and evolution. Astrophysics is primarily an observational subject. Among many other accomplishments, it has provided a convincing picture of the birth, life and death of stars-stellar evolution and the internal structure of stars.

Our knowledge of celestial objects such as the sun, the comets, the stars and the galaxies is mainly based on the analysis of the electromagnetic radiation which arrives to us from these objects.

78.2 Classification of Stars — The Harvard Classification System

Absorption lines were present in the spectra of all the stars (A Continuous spectrum was present too). When the absorption spectra of stars were first studied, it was realized that stars could be classified into several different types, called *spectral classes*. At first, the 'Harvard Classification System' was proposed. In this system, the stars were classified according to the strength of the hydrogen lines in their spectra, and were lettered alphabetically. Class *A* corresponds to the stars having the strongest hydrogen lines, class *B* the next strongest, and so on. It was later realized that the types of spectra varied primarily because of differing temperatures of the stellar atmospheres. So the spectral classes correspond in fact to different surface temperatures. When we now list the spectral types of stars in order of decreasing temperature, they are no longer in alphabetical order. In order of decreasing temperature, the spectral types are *O B A F G K M*. The spectral types can be remembered by the mnemonic : *Oh*, *B*e, *A F*ine *G*irl, *K*iss *M*e. Each class is subdivided into 10 subclasses. For example, the hottest *B* stars are *BO*, followed by *B*1, *B*2, *B*3 ... *B*9. Table 78.1 describes the main characteristics of the spectral classes of stars.

Today we know that the stars are nearly uniform in composition, being composed mainly of hydrogen and helium. The Indian astrophysicist M. N. Saha suggested that the differences in stellar spectra are principally due to different surface temperatures of stars. *M*-type star is the coolest. The gases occur in the form of molecules and molecular band spectra are observed. The spectra of *K* and *M* stars are dominated by molecular bands because they do not have sufficient temperatures to split molecule into atoms. With the increase in temperature, molecules break into neutral atoms and ions. Metals have low ionisation potential. Thus, ionised metal lines are dominant in *F* and *G* type stars. On further increase of temperature, hydrogen is excited and the spectra of *A* type stars show hydrogen lines. At extremely high temperatures, even helium is ionised. Therefore, lines corresponding to ionised helium are present only in *O* — type stars.

TABLE 78.1. Characteristics of the Spectral Classes of Stars

Spectral class	*Temperature range (K)*	*Main characteristics of the spectrum*	*Colour*	*Typical example*
O	30000–50000	Ionized helium lines dominant	Blue	10 Lacertae a Quasar, Zeta Puppis
B	10000–30000	Neutral helium lines dominant, no ionized helium	Blue or Bluish White	Rigel, Spica
A	7500–10000	Hydrogen lines dominant	White	Sirius, Deneb, Vega
F	6000–7500	Ionized calcium lines, Many metal lines (manganese, iron, titanium, strontium)	White	Canopus, Polaris
G	4500–6000	Very large number of metal lines. Strong ionized calcium lines, ionized and neutral iron	Yellow	Sun, Alpha Centauri
K	3500–4500	Large number of neutral metal lines	Red	Arcturus, Aldebaran
M	2000–3500	Band spectra of molecules (particularly of titanium oxide molecule)	Red	Betelgeuse, Antares

78.3 Hertzsprung — Russel Diagram

It is a graphical representation of the absolute magnitude of stars plotted against the spectral class (Fig. 78.1). The *y*-axis then represents the energy output of the star (luminosity) and the *x*-axis its surface temperature. The temperature is plotted from right to left. The majority of stars on such a diagram fall on a band running from the top left to the bottom right of the graph. These are called *main-sequence stars*. At the left and below the main sequence we have a group of stars called *white dwarfs*. The *giants* fall in a cluster above the main sequence and the *supergiants* are above the giants. The H-R diagram forms the basis of the theory of stellar evolution. It has been established that the different groups of stars in the diagram represent different stages in the evolution of a star.

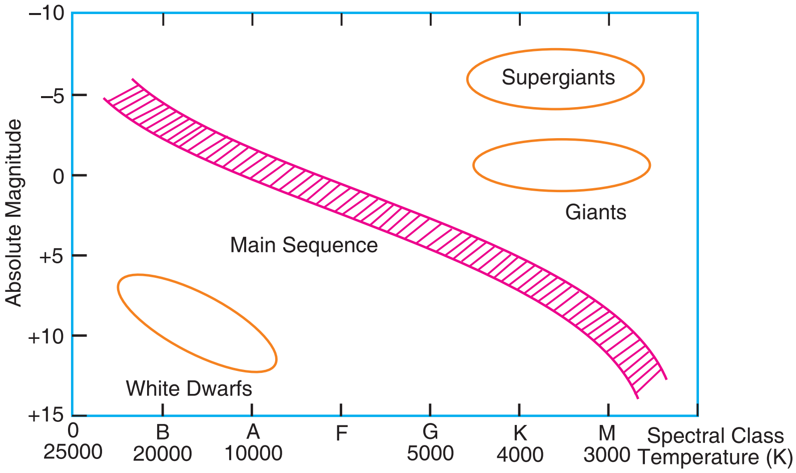


Fig. 78.1

78.4 Luminosity of a Star

The total amount of light energy given out by a star in all directions per unit time is called luminosity L. By Stefan's law,

$$L = 4\pi R^2 \sigma T^4 \qquad \text{...(1)}$$

Here R = radius of the star, T = surface temperature of the star and σ = Stefan constant.

Let d be the distance of the star from us.

$$\left.\begin{array}{l}\text{Energy flux on a unit area}\\ \text{at the observer's eye}\end{array}\right\} = f = \frac{L}{4\pi d^2} \qquad \text{...(2)}$$

Each star is assigned a number called its magnitude, on the basis of its apparent brightness. Since it tells how bright a star *appears*, it is called apparent magnitude (m).

The relation between m and f is

$$m = C - 2.5 \log_{10} f \qquad \text{...(3)}$$

Here, C is a constant which determines the zero of the magnitude scale.

Substituting for f from Eq. (2), we get

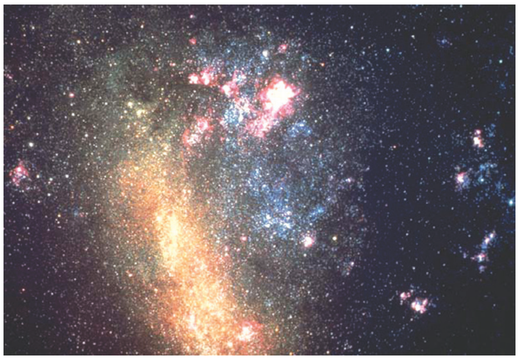
$$m = (C + 2.5 \log_{10} 4\pi) - 2.5 \log_{10} L + 5 \log_{10} d \qquad \text{...(4)}$$

Thus, the apparent magnitude (m) of a star depends on its intrinsic luminosity (L) and its distance (d).

78.5 Stellar Evolution

The life cycle of a one solar mass star such as the sun is expected to follow the following stages: star formation, proto-star stage, pre-main sequence stage, main sequence stage, red giant stage, white dwarf stage.

Birth of a Star



Stars and Stellar Evolution.

The formation of stars is supposed to be the consequence of condensation of the primordial hydrogen under the effect of gravitational attraction. During their random motion in space a few atoms of hydrogen may at some instant of time be converging towards some point. When they are near each other they may exert a small gravitational attraction on each other. If this attraction is sufficiently strong, the atoms would be prevented from flying apart. They would form a pocket of gas. For this to happen the number of atoms converging towards a point must be really very large. Such a pocket of hydrogen atoms forms a beginning of a star. It is called a *protostar*.

Every atom in a protostar is pulled towards the centre of gravity. The protostar as a result shrinks in size. As the atoms fall towards the centre their speed increases. The temperature of the gas increases from about 100 K to about 50,000 K in about a year. At this temperature the hydrogen atoms collide with each other with sufficient violence to dislodge the electrons from the atoms. Instead of the original hydrogen gas there are now two gases, one of protons and the other of electrons. Because of the heat generated the protostar now becomes luminous. But most of the radiation is in the infra-red region. The shrinkage continues with rising temperature.

The shrinkage of the protostar is at first rapid. It then slows down because the density gradually increases and the atoms moving towards each other meet with more resistance. After 10 million years a protostar will have shrunk from its original size of trillions of kilometre to about 2.5 million kilometre in diameter. The internal temperature of the protostar will then be about 10^6 K.

At this temperature, thermonuclear conversion of hydrogen to helium begins. The loss of energy from the outer surface is then just compensated by the total nuclear energy production within the whole interior of the star. When this stage is reached, the star is said to have reached the *main-sequence*. A star is said to be born at this stage. The surface temperature of the new born star is about 3600 K.

The Hertzberg-Russell diagram is the graph obtained by plotting the luminosity of stars (in terms of the luminosity of the sun) against the surface temperature of the stars. The track of a new star of solar mass across the H-R diagram is shown in Fig. 78.2. The age of the star is reckoned from the instant it enters the main sequence.

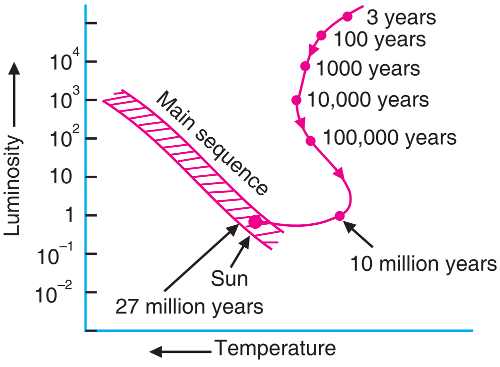


Fig. 78.2

Maturity

During the period the star remains in the main sequence, the release of nuclear energy at the centre balances the inward gravitational force. In the process of hydrogen burning, helium is continually formed as a result of the consumption of hydrogen in its core. This proceeds either by the proton-proton chain reactions or by the carbon-nitrogen cycle. As nuclear energy is released at the centre, the temperature of the star rises. The star becomes more and more luminous. The release of nuclear energy at the centre balances the inward gravitational force. This process of fusion of hydrogen into helium dominates 99% of the life of a star of a solar mass. During this period, the star remains in the main sequence in the H-R diagram.

Ageing stars

The helium formed by fusion collects at the centre. Now there is no burning of hydrogen at the centre. The core contracts. The outer layer expands outwards. As the size of the star increases, the amount of energy radiated per unit area of its surface decreases. As a result, the surface temperature of the star decreases to a value between 3000 K to 4000 K. The temperature drop changes the colour from white to red. The star now leaves the main sequence and heads towards the giants stage. Fig. 78.3 shows the track of a star to the red giant region in the H-R diagram.

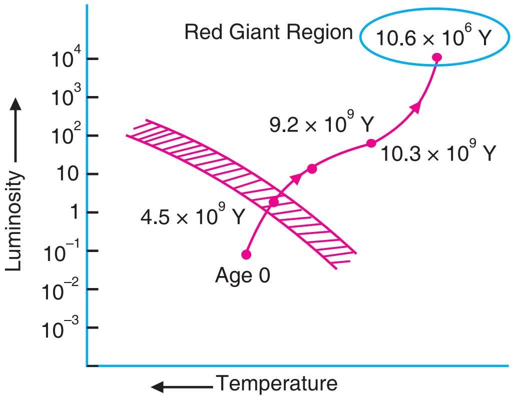


Fig. 78.3

Death of a Star

When the temperature of the helium core reaches 10^8 K, helium nuclei start fusing into carbon through the triple-alpha process. But a star of low mass relatively soon runs out of helium and the core collapse into a *white dwarf*, while the outer regions drift away into space. Larger stars have sufficient helium for the process to continue so that heavier elements, up to iron are formed. But iron is the heaviest element that can be formed with the production of energy. When the helium has all been consumed there is a catastrophic collapse of the whole massive structure, resulting in a supernova explosion. Thereafter the burnt out remnants become neutron stars and black holes.

78.6 White Dwarfs

White Dwarfs.

A small star consisting of elements lighter than iron which has reached the stage where no further nuclear burning is possible is called a white dwarf. Perhaps 10% of the stars in our galaxy are white dwarfs. These are stars in the final stages of their evolution with original masses that were less than about 7 solar masses. After the nuclear reactions that provided it with energy run out of fuel, such a star becomes unstable and eventually throws off its outer layer. The remaining core then cools and contracts gravitationally until its atoms collapse into nuclei and electrons packed closely together. A typical white dwarf has a mass of two-thirds that of the sun but is only about the size of the earth. The density of the white dwarf is quite large. The white dwarf is small and is therefore not easy to detect. The first white dwarf to be detected is the so-called Companion of Sirius.

78.7 Electrons in a White Dwarf Star

Despite the high temperatures found in stellar interiors, the electrons in a white dwarf are quite degenerate because the value of Fermi temperature T_F is so large. When a star burns all its hydrogen fuel, it is made up mainly of a plasma of helium nuclei and electrons. The gravitational forces between the helium nuclei cause the star to collapse until the Pauli pressure of the electrons given by

$$P = \frac{2}{5}\rho\varepsilon_F(0) \qquad \text{...(1)}$$

brings a halt to the collapse.

Here, ρ = Density of the gas, $\varepsilon_F(0)$ = Fermi-energy at absolute zero.

The negative (inward) gravitational pressure is

$$P_G \sim \frac{E_G}{V} \sim \frac{GM^2}{R}\frac{1}{R^3} \qquad \text{...(2)}$$

Here, E_G is the gravitational potential energy, G is the gravitational constant, R is the radius of the star, and $M \sim Nm_{He}$ is the mass of the star made up of N helium atoms each of mass m_{He}.

The outward Pauli pressure is given by

$$\mathrm{P_P} \sim \frac{N}{R^3}\frac{\hbar^2}{m_e}\left(\frac{N}{R^3}\right)^{2/3} \qquad \text{...(3)}$$

Here, m_e is the electron mass and the number of electrons is actually $2N$, twice the number of helium nuclei. Equating P_P and P_G gives the radius of the star.

$$\mathrm{R} \sim \frac{\hbar^2}{Gm_e m_{He}^{5/3} M^{1/3}} \qquad \text{...(4)}$$

We have dropped all numerical constants from this expression to get just an order of magnitude estimate. If we put in numbers we get $R = 7 \times 10^2$ km for $M = M_0$, the solar mass. This means a white dwarf is a very compact object.

Eq. (4) is derived on the basis of nonrelativistic theory. But the electron system of the white dwarf star must be treated relativistically. The relativistic treatment yields the result that, for stellar masses greater than about 1.4 M_0, the Pauli pressure can no longer stop the gravitational collapse. The star continues its infall, a supernova occurs, and ultimately a neutron star or a black hole is formed. This critical mass is known as the *Chandrasekhar limit*.

78.8 Chandrasekhar Limit

The core of a star that has exhausted its nuclear fuel can end up in different ways depending on its mass. In the case of a star like our sun as a result of gravitational compression the core is left with protons, with electrons moving around like the molecules of a gas. This electron gas is able to withstand the gravitational force. As a result a stable equilibrium can be achieved. Such a star is called a white dwarf.

White dwarfs are quite different from all the usual stars. The particles of *classical gases* obey *classical mechanics*. The particles of *quantum gases* obey *quantum mechanics*. Quantum gases help us to understand the behaviour of white dwarfs.

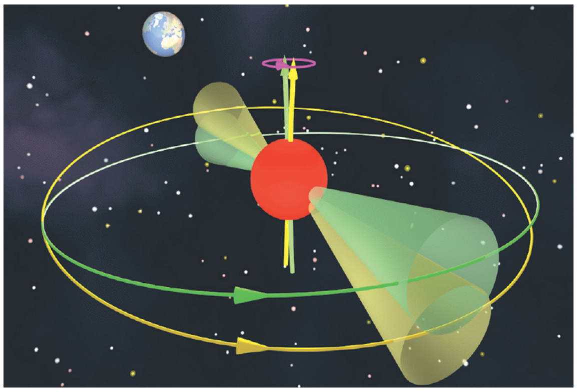
Fowler regarded the core of a white dwarf as a Fermi gas. The pressure in such a gas is of quantum mechanical origin. It is called degeneracy pressure, and has its origins in the Pauli principle. According to Fowler, it was degeneracy pressure which arrested the total collapse of a white dwarf. Fowler applied quantum mechanical ideas to stars.

When the white dwarf gradually cools, at some stage it ceases to be self luminous and ends up as a black dwarf. S. Chandrasekhar discovered theoretically in the 1930s that such a stable condition can occur only for cores which have masses up to 1.4 M_0 where M_0 is the mass of the sun. This is known as the *Chandrasekhar limit*. If the mass of the core exceeds the Chandrasekhar limit, the gravitational compression becomes so large that the electrons are forced into the nuclei. This process is known as inverse beta decay. The star is now composed only of neutrons. These neutrons can withstand the gravitational compression. We now have a neutron star.

78.9 Neutron Stars

Let us now discuss the fate of the core of a star that explodes as a supernova. As iron fills the core of a massive star, the temperatures are so high that the iron nuclei begin to break apart into smaller units like alpha particles (helium nuclei). The pressure is no longer high enough to counteract gravity and the core collapses. As the density increases, the electrons are squeezed into the nuclei and react with the protons there to produce neutrons and neutrinos.

$$p^+ + e^- \rightarrow n + \nu$$



Neutron Star.

The neutrinos escape. A gas composed mainly of neutrons is left behind in the dense core as the outer layers explode as a supernova. Now the core has a very high density. At this density a condition called *neutron degeneracy*, in which the neutrons cannot be packed any more tightly, appears. The pressure caused by neutron degeneracy balances the gravitational force that tends to collapse the core. As a result, the core reaches equilibrium as a *neutron star*.

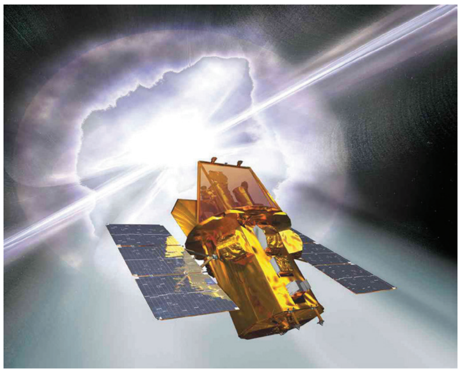
Typically neutron stars have a radius of about 10 km, although in mass they are comparable with the sun. The chunk of their matter the size of a small sugar cube contains about 10^8 tons of neutrons. Neutron stars have masses between 1.4 and ~ 3 M_{sun}.

Before it collapses, the core has only a weak magnetic field. But as the core collapses, the magnetic field is concentrated, and grows stronger as a result. By the time the core shrinks to neutron star size, it has an extremely powerful magnetic field.

Neutron stars were postulated in the 1930s and many astronomical observations confirm their existence. Neutron stars are believed to be pulsars if they rotate sufficiently rapidly and have sufficiently strong magnetic fields.

Pulsars. First discovered in 1968, a pulsar is believed to be a rotating neutron star. They emit radiation from relatively a small spot on the star. This situation produces an effect rather like the beam of a lighthouse. The distant observer receives a pulse of radiation as the beam sweeps across his position on each rotation of the star.

A pulsar gives off pulses of radio waves. A very few objects (example : Crab and the Vela pulsars) are known to pulse γ-rays, X-rays, or visible light. The periods of known pulsars range between 33 milliseconds and 3.75 seconds. Pulse durations range from 2 to about 150 milliseconds with longer period pulsars generally having a longer pulse duration.



New Born Black Holes.

78.10 Black Holes

Consider stars with very large mass, say 5 to 10 M_{sun}. For such a star, the contraction cannot be arrested either at the white dwarf stage or at the stage of the neutron star. A star may continue to collapse beyond the neutron star stage. When the radius of the star is of the order of 15 kilometre, relativity predicts a most extraordinary phenomenon. According to the theory of relativity, a ray of light should possess mass and hence be subject to gravity. A ray of light emitted by a star will therefore be pulled back by the star's gravity. If the star is large in size, gravity will not be strong at its surface and a ray of light will be able to escape from the star. If the star shrinks to a size of about 15 kilometre radius, the force of gravity at its surface will be billions of times stronger than the force of gravity at the surface of the sun. A ray of light trying to leave the star will therefore be pulled back and it cannot escape from the star. And when light cannot escape, nothing else can escape from it. The star then becomes invisible. It becomes a *black hole* in space.

The contraction continues inside a black hole. There is no force in nature which is strong enough to halt this contraction. The volume of the star must become smaller and smaller. The mass of the star however remains constant. Ultimately trillions and trillions of quintals of matter will have to be packed in a size less than a pinhead.

Observational evidence of objects thought to be black holes comes from their effect on surrounding matter. Thus, if a black hole forms a binary system with another star it will attract and capture matter from this star. The material leaving the star first forms a rotating disc around the black hole, in which the matter becomes compressed. The disc is heated to such an extent that it emits X-rays. In the constellation Cygnus there is an X-ray source, Cygnus X-1, which consists of a supergiant revolving around a small invisible companion with a mass ten times that of the sun. The companion is thought to be a black hole.

The Basic Physics of Black Holes

Consider an escape velocity such that, when an object leaves with just the escape velocity, it will have zero velocity out at "infinity". There, its total energy is

$$TE = (KE + PE) = (1/2)\, mv^2 - GmM/R = 0.$$

Since total energy must be conserved, at the moment of launch we require

$$\begin{aligned} TE &= 0 = (1/2)\, mV^2_{\text{esc}} - GmM/R \\ (1/2)\, V^2_{\text{esc}} &= GM/R \\ V_{\text{esc}} &= (2GM/R)^{1/2} \end{aligned}$$

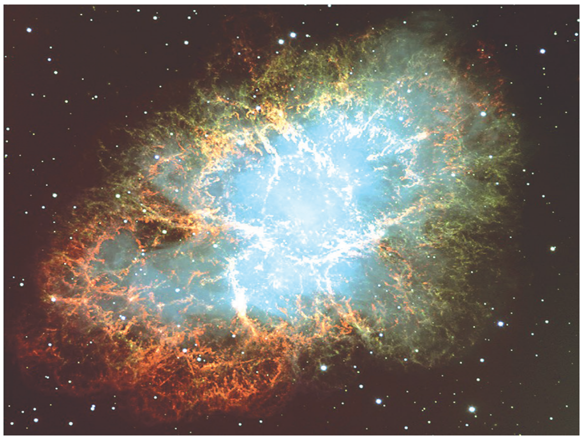
Now no object can travel faster than the speed of light, and so the maximum escape velocity is c. Then the equation for the black-hole radius is

$$R = 2GM/c^2.$$

This critical radius is called the *Schwarzschild radius*.

78.11 Supernova Explosion

When a massive star starts rapidly collapsing, the interior gets heavily compressed and therefore also very hot. The core becomes converted into neutrons. If the compression is sudden and the heat generated is very high, then a violent explosion called *supernova explosion* takes place. It tears off the outer layers of the star and hurls them into space. What is left after a supernova explosion may be a remnant whose mass is greater than 1.4 M_{sun}. The Crab Nebula is the nearest *supernova remnant*.



Supernova Explosion.

78.12 Photon Diffusion Time

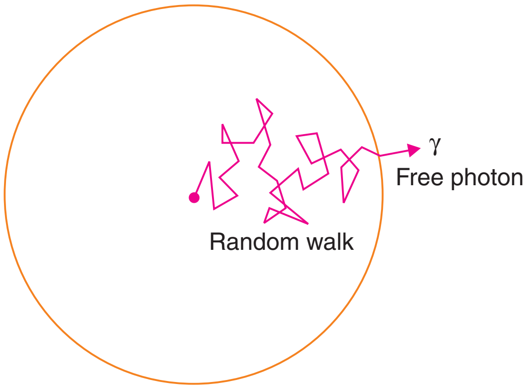


Fig. 78.4

Photons are produced in the core of the star due to thermonuclear reactions. Photons move from the core towards outer surface of the star. The photons get scattered large number of times by free electrons before reaching the surface. This results in an apparent dance like motion of the photons inside the stars like the sun (Fig. 78.4).

This dance like motion of the photons is called *random walk*. The photon travels along a *zig-zag* path from the core of the sun to the sun's surface. Finally the photon escapes from the sun's surface.

During their random walk, the photons which are X-ray photons at the core become optical photons by the time they reach the surface of the star by *diffusion*.

The amount energy radiated out per second from the surface of a star is its *luminosity L*.

The **photon diffusion** time is defined as the ratio of the total radiant energy E contained in the volume of the star to the luminosity L of the star. *i.e.,*

$$\tau = \frac{E}{L} \qquad \text{... (1)}$$

The total amount of light energy given out by a star in all directions per unit time is called luminosity L. By Stefan's law,

$$L = 4\pi R^2 \sigma T^4 \qquad \text{... (2)}$$

Here R = radius of the star, T = surface temperature of the star and σ = Stefan constant.

The energy density of the radiation which is a function of temperature $T(r)$ is defined as

$$U = \frac{4\sigma}{c} T^4(r) \qquad \text{... (3)}$$

Here, c = velocity of light.

According to the linear star model, the relation between the core temperature T_c and the varying temperature $T(r)$ is given by

$$T(r) = T_c \left[1 - \frac{r}{R}\right] \qquad \text{... (4)}$$

$$\therefore \quad U = \frac{4\sigma}{c} T_c^4 \left(1 - \frac{r}{R}\right)^4 \qquad \text{... (5)}$$

The total radiant energy = $E = \int_0^R U\, dV = \int_0^R \frac{4\sigma}{c} T_c^4 \left(1 - \frac{r}{R}\right)^4 dV$

But $dV = 4\,\pi r^2\, dr$.

$$E = \int_0^R \frac{4\sigma T_c^4}{c} \left(1 - \frac{r}{R}\right)^4 4\pi r^2\, dr = \frac{16\pi\sigma T_c^4}{c} \int_0^R \left(1 - \frac{r}{R}\right)^4 r^2 dr$$

$$\therefore \quad E = \frac{1}{105}\left(\frac{16\,\pi\sigma}{c}\right) R^3 T_c^4 \qquad \text{... (6)}$$

For the sun, $E = 1.4 \times 10^{39}$ J and $T_c = 15 \times 10^6$ K.

Substituting E and L from Eqs. (6) and (2) in Eq. (1), we get

$$\tau = \frac{\frac{1}{105}\left(\frac{16\,\pi\sigma}{c}\right)R^3T_c^4}{4\pi R^2\sigma T^4}$$

$$\therefore \qquad \tau = \frac{4R}{105c}\left(\frac{T_c}{T}\right)^4 \qquad \dots (7)$$

For the sun, $\tau = 3.46 \times 10^{12} s = 1.2 \times 10^5$ years.

Mass - Luminosity relation for a star

From Eq. (1), $L = \dfrac{E}{\tau}$ ... (8)

From Eq. (6), $E \propto R^3T_c^4$ ... (9)

It can be shown that $T_c \propto \dfrac{M}{R}$ ... (10)

Here, M is the mass of the star.

$$\therefore \qquad E \propto R^3\left(\frac{M}{R}\right)^4 \propto \frac{M^4}{R} \qquad \dots (11)$$

The mean free path (s) of the diffusing photon is defined as the distance travelled by the photon between its two successive scatterings.

The photon diffusion time $\tau \propto \dfrac{R^2}{s}$ ... (12)

The mean free path (s) is inversely proportional to the mean photon density. *i.e.*,

$$s \propto \frac{1}{\rho} \propto \frac{V}{M} \propto \frac{R^3}{M}$$

$$\therefore \qquad \tau \propto \frac{R^2M}{R^3} \times \frac{M}{R} \qquad \dots (13)$$

Substituting for E and τ from Eqs. (11) and (13) in Eq. (8), we get

$$L \propto \frac{M^4}{R} \times \frac{R}{M}$$

$$\therefore \qquad L \propto M^3 \qquad \dots (14)$$

The life time of a star is defined as the ratio between the mass and the luminosity of a star.

$$\text{Life time} \propto \frac{M}{L} \propto \frac{M}{M^3}$$

$$\therefore \text{ Life time of a star} \propto \frac{1}{M^2}$$

The life time of a star is inversely proportional to the square of its mass.

78.13 Gravitational Potential Energy of a Star

The formation of a star is due to the condensation of the primordial hydrogen under the effect of gravitational attraction. Consider a star of mass M and radius R being formed by a slow gravitational contraction of hydrogen gas (Fig. 78.5).

Initially the gas particles are at an infinite distance. Let $dM(r)$ be the mass of a cloud of spherical shell of thickness dr, internal radius r and density $\rho(r)$ over a sphere of mass $M(r)$ and radius r.

$$dM(r) = 4\pi r^2\, \rho(r)\, dr \qquad \text{... (1)}$$

The work done to bring this mass $dM(r)$ and add it to the sphere of mass $M(r)$ is given by the gravitational potential energy,

$$dU = -\frac{G\,M(r)\,dM(r)}{r} \qquad \text{... (2)}$$

The total gravitational energy required to form the entire star is given by,

$$U = -\int_0^R \frac{G\,M(r)\,dM(r)}{r} \qquad \text{... (3)}$$

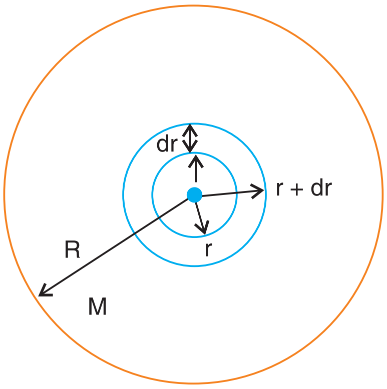


Fig. 78.5

(*i*) According to **constant density model of a star**, density is constant throughout the interior of a star.

$$\therefore \qquad M(r) = \int_0^r 4\pi r^2\, \rho\, dr = \frac{4}{3}\pi r^3 \rho \qquad \text{[from Eq. (1)]}$$

Substituting for $dM(r)$ and $M(r)$ in Eq. (3), we have

$$U = -G\int_0^R \frac{\left(\frac{4\pi r^3}{3}\rho\right)(4\pi r^2\,\rho dr)}{r} = -G\,\frac{16\pi^2\rho^2}{3}\int_0^R r^4\,dr = -G\,\frac{16\pi^2\rho^2}{15}R^5 = -G\,\frac{3}{5R}\left(\frac{4}{3}\pi R^3\rho\right)^2$$

$M = \frac{4}{3}\pi R^3\rho$, where ρ is the average density.

$$\therefore \qquad U = -\frac{3}{5}\frac{GM^2}{R}. \qquad \text{... (4)}$$

Eq. (4) gives the gravitational potential energy of a star as per constant density model of a star.

(*ii*) According to the *linear star model*, density $\rho(r)$ decreases from maximum ρ_c to 0.

$$\rho(r) = \rho_c\left(1 - \frac{r}{R}\right) \qquad \text{... (5)}$$

Substituting this value of $\rho(r)$ in Eq. (1), we have

$$dM(r) = 4\pi r^2\, \rho_c\left(1 - \frac{r}{R}\right)dr \qquad \text{... (6)}$$

$$M(r) = 4\pi\,\rho_c\int_0^r\left(r^2 - \frac{r^3}{R}\right)dr = 4\pi\,\rho_c\left[\frac{r^3}{3} - \frac{r^4}{4R}\right] = \frac{4}{3}\pi r^3\,\rho_c\left(1 - \frac{3r}{4R}\right)$$

The total mass M of the star (when $r = R$) is given by

$$M = \frac{\pi R^3}{3}\rho_c$$

or

$$\rho_c = \frac{3M}{\pi R^3} \qquad \text{... (7)}$$

But average density is given by $\rho = \dfrac{3M}{4\pi R^3}$

$$\therefore \qquad \rho_c = 4\rho \qquad \text{... (8)}$$

Substituting for $dM(r)$ and $M(r)$ in Eq. (3),

$$\text{U} = -G\int_0^R \frac{\frac{4\pi}{3}\rho_c\, r^3\left(1-\frac{3r}{R}\right)\times 4\pi r^2\, \rho_c\left(1-\frac{r}{R}\right)dr}{r} = -\frac{G(4\pi\rho_c)^2}{3}\int_0^R r^4\left(1-\frac{3r}{4R}\right)\left(1-\frac{r}{R}\right)dr$$

Substituting for ρ_c,

$$U = -\frac{G}{3}\left(\frac{4\pi\, 3M}{\pi R^3}\right)^2\int_0^R r^4\left(1-\frac{3r}{4R}\right)\left(1-\frac{r}{R}\right)dr = -\frac{144}{3}\frac{GM^2}{R^6}\left[\frac{312\, R^5}{20160}\right] = -\frac{26}{21}\left(\frac{3}{5}\frac{GM^2}{R}\right)$$

But $\dfrac{26}{21} \cong \dfrac{5}{4}$.

$$\therefore \qquad U = -\frac{3}{4}\frac{GM^2}{R}. \qquad \text{... (9)}$$

78.14 Internal Temperature of a Star

A contracting star radiates energy just like a black body and loses energy obeying Stefan-Boltzmann law ($E = \sigma T^4$). To uphold the law of conservation of energy, half the loss in potential energy U of the star is gained by the star itself in the form of kinetic energy K and the remaining half of the energy is lost by radiation.

$$\therefore \qquad K = U/2 \qquad \text{... (1)}$$

The interior of a star behaves like an ideal gas. According to kinetic theory of gases, the kinetic energy of the interior gas is given by

$$K = \frac{3}{2} Nk\, T_{av} \qquad \text{... (2)}$$

Here, N is the number of gas particles, k is the Boltzmann constant and T_{av} is the average temperature of the entire stellar matter.

Let m = mass of each gas particle and M = mass of entire star. Then,

N = total number of particles contained in the star = M/m.

$$\therefore \qquad K = \frac{3}{2}\left(\frac{M}{m}\right)k\, T_{av} \qquad \text{... (3)}$$

According to *Virial theorem*, the kinetic energy K is equal to the total thermal energy of the star and is given by

$$K = \frac{3}{8}\frac{GM^2}{R} \qquad \text{... (4)}$$

Comparing Eqs. (3) and (4),

$$T_{av} = \frac{GMm}{4kR}$$

If μ is the mean molecular weight in terms of atomic mass of hydorgen $m_H = 1.67 \times 10^{-27}$ kg, then $m = \mu m_{\text{H}}$.

$$\therefore \qquad T_{av} = \frac{GM\mu m_H}{4kR} \qquad \text{... (5)}$$

This is the first equation of stellar structure.

If all the hydrogen atoms in a star (sun) are considered to be completely ionized, then $\mu = 0.5$. Sun is a main sequence star. Its properties like surface temperature, mass, radius, etc., are well known.

The average temperature of any star of similar composition can be determined in terms of solar parameters using the relation

$$T_{av} = \left(\frac{M}{M_{sun}}\right)\left(\frac{R_{sun}}{R}\right)(T_{sun})_{av} \qquad \text{... (6)}$$

Let $T(r)$ be the internal temperature of the sun at a distance r from the centre. Let T_c = core temperature.

According to the linear star model $T(r) = T_c\left[1 - \frac{r}{R}\right]$ ... (7)

Here, $T_c = T_{av}$ at the centre of the star and $r = R$ at the sruface.

78.15 Internal Pressure of a Star

Consider the given star to be made up of large number of concentric shells. Each shell is in *hydrostatic equilibrium*. Consider a shell of radius r and thickness dr. Let $\rho(r)$ be the density of stellar matter in the shell.

Mass of the shell, $dM(r) = 4\pi r^2 \rho(r)\, dr$

or $$\frac{dM(r)}{dr} = 4\pi r^2 \rho(r) \qquad \text{... (1)}$$

If the stellar mass distribution is constant, then Eq. (1) represents conservation of stellar mass. Therefore, the outward pressure on the shell from within the star is equal to the gravitational force acting on the shell towards the centre of the star.

The gravitational force between the shell of mass $dM(r)$ and the stellar sphere of mass $M(r)$ with radius r is given by

$$F_{gr} = \frac{-G\,M(r)\,dM(r)}{r^2} = \frac{-GM(r)\,4\pi r^2\,\rho(r)\,dr}{r^2} \qquad \text{... (2)}$$

Let $P(r)$ = outward gas pressure on the shell from within the star and

$P(r + dr)$ = inward gas pressure on the sheel from above it (Fig. 78.6).

Net outward pressure acting on the shell is given by $dP = P(r + dr) - P(r)$

The total outward force acting on the shell to balance it against the inward gravitational force is given by

$$F_{pr} = 4\pi r^2\, dP \qquad \text{... (3)}$$

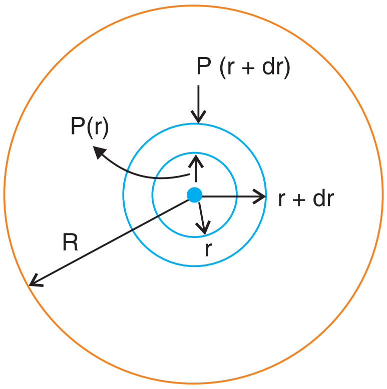


Fig. 78.6

For *hydrostatic balance*,

$$F_{pr} = F_{gr}$$

$$\therefore \quad \frac{dP}{dr} = \frac{-G\,M(r)\,\rho(r)}{r^2} \qquad \text{... (4)}$$

Eq. (4) is known as the equation of *hydrostatic equilibrium* and is the *second equation* of *stellar structure*. All stable main sequence stars obey this equation.

If the stellar mass is considered as an ideal gas, the ideal gas equation is

$$PV = RT = N_A kT.$$

Here, N_A = Avogadro number.

$$P = \frac{N_A}{V} kT = nkT$$

Here, $n = \frac{N_A}{V}$ is the molecular density.

Let ρ_{av} = average stellar density and μ = mean molecular weight in the star.

Molecular density,

$$n = \frac{\rho_{av}}{\mu\, m_H} = \frac{\frac{M}{\left(\frac{4}{3}\pi R^3\right)}}{\mu\, m_H} \qquad \text{... (5)}$$

The average temperature of a star is given by

$$T_{av} = \frac{GM\mu m_H}{4kR} \qquad \text{... (6)}$$

If the average pressure and temperature of the star are considered, then

$$P_{av} = nkT_{av} = \frac{\rho_{av}}{\mu m_H} kT_{av} \qquad \text{... (7)}$$

Substituting for n and T_{av}, we have

$$P_{av} = \frac{3}{16\pi}\frac{GM^2}{R^4} \qquad \text{... (8)}$$

Eq. (7) gives the average internal pressure of a star.

The average pressure for any star can be determined in terms of solar parameters using the relation

$$P_{av} = \left(\frac{M}{M_{sun}}\right)^2 \left(\frac{R_{sun}}{R}\right)^4 (P_{sun})_{av} \qquad \text{... (9)}$$

From Eq. (7), the internal (or core) pressure is given by

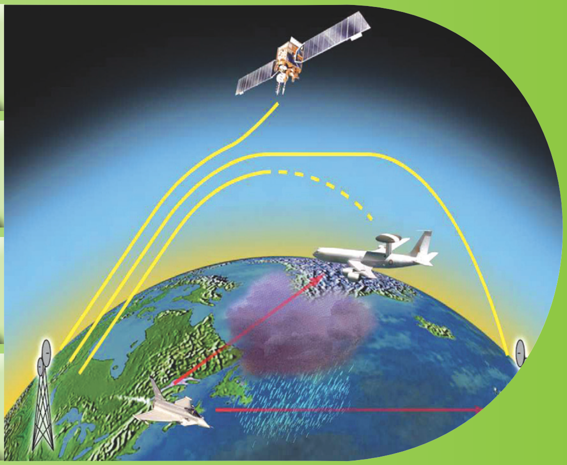
$$P_c = \frac{\rho_c}{\mu m_H} kT_c \qquad \text{... (10)}$$

The internal pressure of the Sun at a distance r from its centre is given by

$$P(r) = P_c\left[1 - \frac{r}{R}\right]. \qquad \text{... (11)}$$

CHAPTER

79



NANOTECHNOLOGY

AT A GLANCE

79.1	Introduction	79.2	Nanoparticles
79.3	Metal Nanoclusters	79.4	Semiconducting Nanoparticles
79.5	Rare Gas and Molecular Clusters	79.6	Carbon Clusters
79.7	Carbon Nanotubes	79.8	Applications of Carbon Nanotubes
79.9	Quantum Nanostructures	79.10	Preparation of Quantum Nanostructures
79.11	Size and Dimensionality Effects	79.12	Applications of Quantum Nanostructures
79.13	Photolithography	79.14	Scanning Tunnelling Microscopy (STM)
79.15	Atomic Force Microscopy (AFM)		

79.1 Introduction

The prefix *nano* in the word *nanotechnology* means a billionth (1×10^{-9}). Nanotechnology deals with various structures of matter having dimensions of the order of a billionth of a meter. A nanometer is 10^{-9} m. Nanoparticles are generally considered to be a number of atoms or molecules bonded together with a radius of < 100nm. Generally, physical properties of materials can be characterized by some critical length, a thermal diffusion length, or a scattering length, for example. The electrical conductivity of a metal is strongly determined by the distance that the electrons travel between collisions with the vibrating atoms or impurities of the solid. This distance is called the *mean free path* or the *scattering length.* If the sizes of particles are less than these characteristic lengths, it is possible that new physics or chemistry may occur. *A working definition of a nanoparticle is an aggregate of atoms between 1 and 100 nm viewed as a subdivision of a bulk material, and of dimension less than the characteristic length of some phenomena.*

Materials with at least one of the dimensions measuring less than 100 nm are known as *nanomaterials. We define nanomaterials as those which have a characteristic length scale within about 100 nm.*

A particle diameter, grain size, layer thickness, or width of a conducting line on a device are some examples of length scale.

Example. *The size of the Quantum dot is 1–5 nm.*

Special Features of Nanomaterials

Conventional materials have grains varying in size anywhere from hundreds of microns (μm) to millimeters (mm). A nanocrystalline material has grains on the order of 1-100 nm. The properties of bulk materials are mostly retained till the reduction of their dimensions to the micrometer range. But materials in the nanometre scale show remarkably new properties.

Example. *1. A nanocrystal of gold serves as an excellent low temperature catalyst whereas the bulk gold does not.*

2. Bulk semiconductors transform into insulators if their dimensions shrink to a couple of nanometres.

3. The melting temperature of crystals in the nanometre scale is very low, lower by 1000°C in some cases. The lattice constants are, too, reduced.

4. A nanoscale wire or circuit component does not necessarily obey the Ohm's law of electricity.

79.2 Nanoparticles

A group of 10^6 or less number of atoms or molecules bonded together in a cluster with the radius of about 100 nm or less forms a nanoparticle. The metal oxide nanoparticles of different sizes in stained-glass windows produce different beautiful colours because a particle scatters only that wavelength which compares with its size. The understanding of structure and properties of nanoparticles is crucial for the interpretation of the characteristic phenomena exhibited by them. A brief description of some well-studied nanoparticles is given below.

79.3 Metal Nanoclusters

79.3.1. Magic Numbers

Experimental set-up to form clusters of metal atoms. Fig. 79.1 shows the apparatus to make metal nanoparticles by laser induced evaporation of atoms from the surface of a metal.

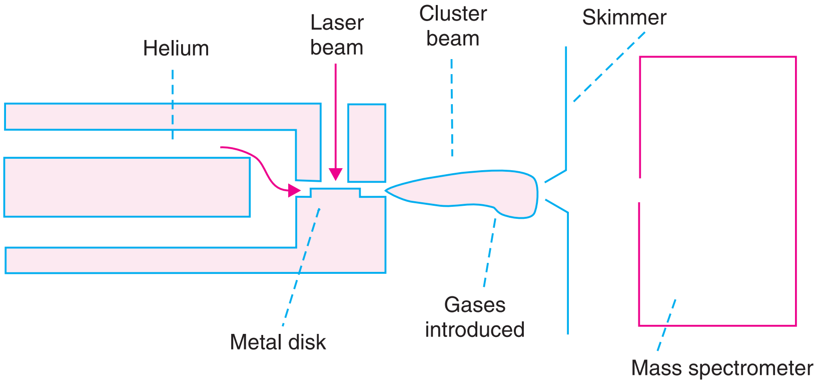


Fig. 79.1

A high-intensity laser beam is incident on a metal rod, causing evaporation of atoms from the surface of the metal. The atoms are then swept away by a burst of helium and passed through an orifice into a vacuum. In the evacuated chamber, the expansion of the gas causes cooling and formation of clusters of the metal atoms. These clusters are then ionized by UV radiation and passed into a mass spectrometer. The mass spectrometer measures their mass : charge ratio.

Mass spectrum of Pb clusters: Fig. 79.2 shows the mass spectrum data of lead clusters formed in the experiment. Here, the number of ions (counts) is plotted as a function of the number of atoms in the cluster. (Usually mass spectra data are plotted as counts versus mass over charge.) The data show that clusters of 7 and 10 atoms are more likely than other clusters which means that these clusters are more stable than clusters of other sizes.

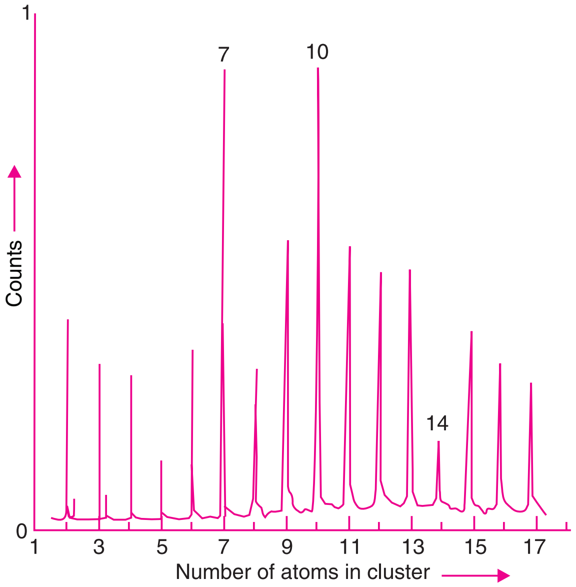


Fig. 79.2

Ionization Potential: Fig. 79.3 (*a*) is a plot of the ionization potential of atoms as a function of their atomic number Z. The maximum ionization potentials occur for the rare-gas atoms ^{2}He, ^{10}Ne and ^{18}Ar because their outermost *s* and *p* orbitals are filled. More energy is required to remove electrons from filled orbitals than from unfilled orbitals.

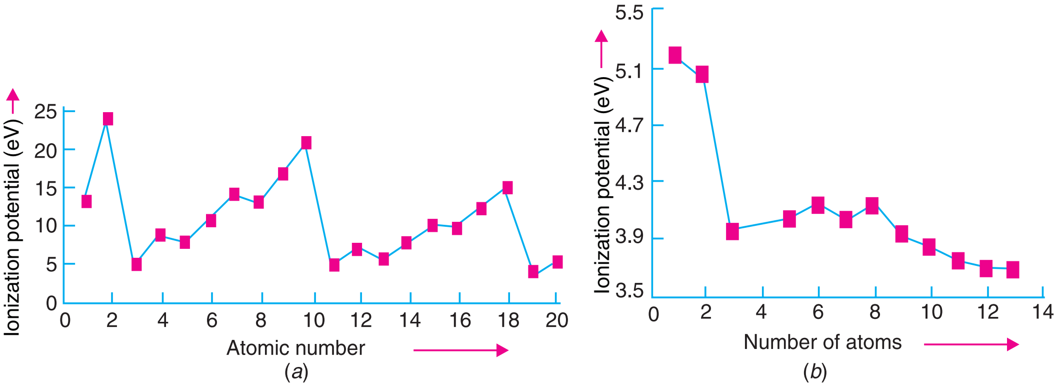


Fig. 79.3

Fig. 79.3 (*b*) shows the ionization potential of sodium clusters as a function of the number of atoms in a cluster. Peaks are observed at clusters having two and eight atoms. These numbers are referred to as *electronic magic numbers*. Their existence suggests that clusters can be viewed as superatoms. This result motivated the development of the jellium model of clusters. In the case of larger clusters, stability is determined by structure and the magic numbers are referred to as *structural magic numbers*.

79.3.2. Theoretical Modeling of Nanoparticles

There exist two well-known theoretical models to calculate the properties of clusters:

(*i*) the jellium model and (*ii*) the density functional model (based on the molecular orbital theory).

(*i*) The jellium model. The jellium model envisions a cluster of atoms as a large atom. The positive nuclear charge of each atom of the cluster is assumed to be uniformly distributed over a sphere the size of the cluster. A spherically symmetric potential well is used to represent the potential describing the interaction of the electron with the positive spherical charge distribution. Thus the energy levels can be obtained by solving the Schrödinger equation for this system in a fashion analogous to that for the hydrogen atom.

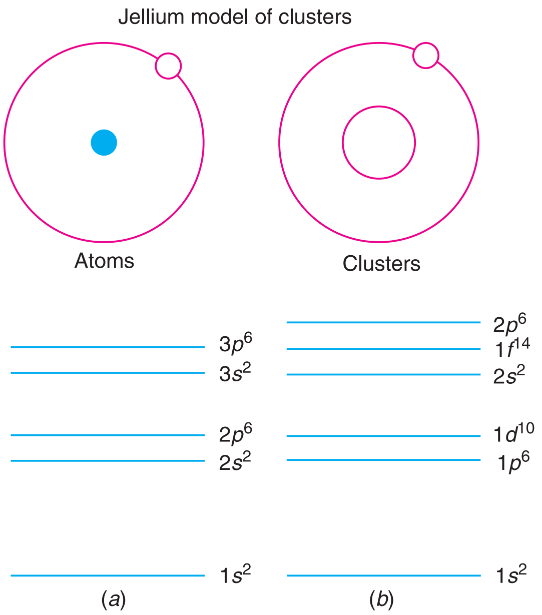


Fig. 79.4. A comparison of the engery levels of the hydrogen atom and those of the jellium model of a cluster.

Fig. 79.4 compares the energy level scheme for the hydrogen atom and the energy-level scheme for a spherical positive-charge distribution. The superscripts refer to the number of electrons that fill a particular energy level. The electronic magic number corresponds to the total number of electrons on the superatom when the top level is filled. Notice that the order of the levels in the jellium model is different from that of the hydrogen atom. In this model the magic numbers correspond to those clusters having a size in which all the energy levels are filled. The electronic magic numbers of the atoms are 2, 10, 18, and 36 for He, Ne, Ar, and Kr, respectively (the Kr energy levels are not shown on the figure) and 2, 18 and 40 for the clusters.

(*ii*) The density functional model

This model treats the clusters as molecules and uses existing molecular orbital theories such as density functional theory to calculate their properties. This approach can be used to calculate the actual geometric and electronic structure of small metal clusters.

The simplest of clusters is a molecule such as H_2^+ because the Schrödinger equation can be solved exactly for a hydrogen atom. The ground state wavefunction of the hydrogen atom is expressed as

$$\psi(1s) = A \exp\left(-\frac{r}{\rho}\right) \quad \text{...(1)}$$

Here, r is the distance of the electron from the nucleus and ρ is the radius of the first Bohr orbit.

In the case of a molecule such as the H_2^+ ion, molecular orbital theory assumes that the wavefunction of the electron around the two H nuclei can be described as a linear combination of the wavefunction of the isolated H atoms. Thus the wavefunction of the electrons in the ground state will have the form,

$$\psi = a\psi(1)_{1s} + a\psi(2)_{1s} \quad \text{...(2)}$$

The Schrödinger equation for the molecular ion is

$$\left[\left(\frac{-\hbar^2}{2m}\right)\nabla^2 - \frac{e^2}{4\pi\varepsilon_0 r_a} - \frac{e^2}{4\pi\varepsilon_0 r_b}\right]\psi = E\psi \qquad ...(3)$$

Here, r_a and r_b are the distances of the electron from the two nuclei '*a*' and '*b*'.

The last two terms in the brackets are the electrostatic attraction of the electron to the two positive nuclei.

For the hydrogen molecule, which has two electrons, a term for the electrostatic repulsion of the two electrons would be added. The Schrödinger equation is solved to calculate the actual electronic and geometric structures of small clusters.

When there are many atoms in the molecule and many electrons, the problem becomes complex, and many approximations are used to obtain the solution. Density functional theory represents one approximation. With the development of large fast computer capability and new theoretical approaches, it is possible using molecular orbital theory to determine the geometric and electronic structures of large molecules with a high degree of accuracy. The calculations can find the structure with the lowest energy, which will be the equilibrium geometry. These molecular orbital methods with some modification have been applied to metal nanoparticles.

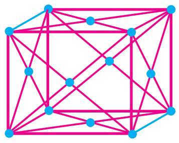


Fig. 79.5

79.3.3. Geometric Structure

Generally the crystal structure of large nanoparticles is the same as the bulk structure with somewhat different lattice parameters. X-ray diffraction studies of 80-nm aluminum particles have shown that it has the face-centered cubic (FCC) unit cell (Fig. 79.5), which is the structure of the unit cell of bulk aluminum.

But small particles having diameters of < 5 nm may have different structures. For example, it has been shown that 3-5 nm gold particles have an icosahedral structure rather than the bulk FCC structure.

Indium clusters undergo a change of structure when the size is smaller than 5.5 nm. Above 6.5 nm, a diameter corresponding to about 6000 atoms, the clusters have a face-centered tetragonal structure with a *c*/*a* ratio of 1.075. In a tetragonal unit cell the edges of the cell are perpendicular, the long axis is denoted by *c*, and the two short axes by *a*. Below ~ 6.5 nm the *c*/*a* ratio begins to decrease. At 5 nm *c*/*a* = 1, meaning that the structure is face-centered cubic. Fig. 79.6 is a plot of *c*/*a* versus the diameter of indium nanoparticles. A different structure can result in a change in many properties.

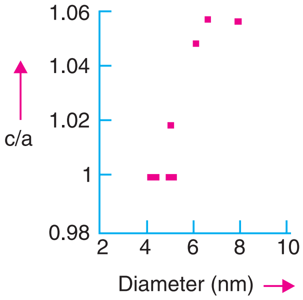
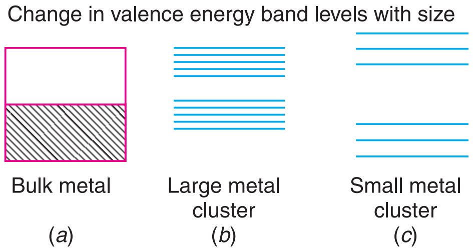


Fig. 79.6



(*a*) valence band of *bulk metal;*
(*b*) *large metal cluster* of 100 atoms showing opening of a band gap;
(*c*) *small metal cluster* containing three atoms.

Fig. 79.7

79.3.4. Electronic Structure

When a metal particle having bulk properties is reduced in size to a few hundred atoms, the density of states in the conduction band changes dramatically. Fig. 79.7 shows the changes in the electronic structure during the transition of a bulk metal to a large cluster, and then down to a small cluster of less than 15 atoms. The figure illustrates how energy levels of a metal change when the number of atoms of the material is reduced.

(*i*) At a size when a particle has a few hundred atoms or less, the continuous density of states changes over to a set of discrete energy levels. If the spacing between discrete energy levels turns out to be greater than the thermal energy k_BT, a gap opens up.

(*ii*) The small cluster is analogous to a molecule having discrete energy levels with bonding and antibonding orbitals.

Eventually a size is reached where the surfaces of the particles are separated by distances which are in the order of the wavelengths of the electrons. In this situation, the energy levels can be modeled by the quantum-mechanical treatment of a particle in a box. This is referred to as the *quantum size effect*.

The emergence of new electronic properties can be understood in terms of the Heisenberg uncertainty principle, which states that the more an electron is spatially confined the broader will be its range of momentum. The average energy will not be determined so much by the chemical nature of the atoms, but mainly by the dimension of the particle.

Reactivity: Since the electronic structure of nanoparticles depends on the size of the particle, the ability of the cluster to react with other species should depend on cluster size. This has important implications for the design of catalytic agents.

79.3.5. Magnetic Clusters

In this section we discuss the magnetic properties of nanoclusters of metal atoms that have a net magnet moment. In a cluster the magnetic moment of each atom will interact with the moments of the other atoms, and can force all the moments to align in one direction with respect to some symmetry axis of the cluster. The cluster will have a net moment, and is said to be magnetized. The magnetic moment of such clusters is measured by Stern-Gerlach experiment.

Clusters made up of nonmagnetic atoms can have a net magnetic moment. For example, clusters of rhenium show a pronounced increase in their magnetic moment when they contain less than 20 atoms. Fig. 79.8 is a plot of the magnetic moment versus the size of the rhenium cluster. The magnetic moment is large when n is less than 15.

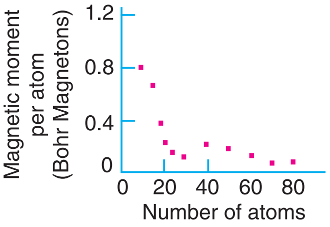


Fig. 79.8

79.4 Semiconducting Nanoparticles

79.4.1. Optical Properties

In this section we discuss nanoparticles made of the elements, which are normal constituents of semiconductors. Nanoparticles made of cadmium, germanium, or silicon are not themselves semiconductors.

Synthesis of Si_n nanoparticles. A nanoparticle of Si_n can be made by laser evaporation of a Si substrate in the region of a helium gas pulse. The beam of neutral clusters is photolyzed by a UV laser producing ionized clusters whose mass to charge ratio is then measured in a mass spectrometer.

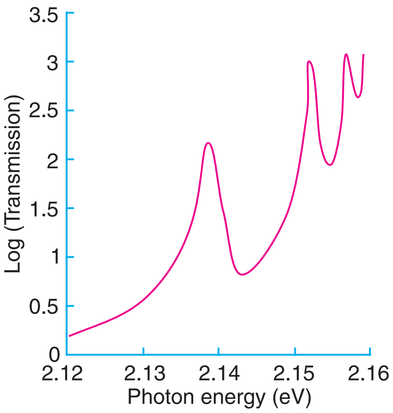


Fig. 79.9

Optical absorption spectrum: In a bulk semiconductor a bound electron-hole pair, called an *exciton*, can be produced by a photon having an energy greater than that of the band gap of the material. The photon excites an electron from the filled band to the unfilled band above. The result is a hole in the valence band. Because of the Coulomb attraction between the positive hole and the negative electron, a bound pair, called an *exciton*, is formed.

The exciton can be modeled as a hydrogen-like atom. Light-induced transitions between these hydrogen-like energy levels produce a series of optical absorptions. Fig. 79.9 shows the optical absorption spectrum of excitons in cuprous oxide (Cu_2O).

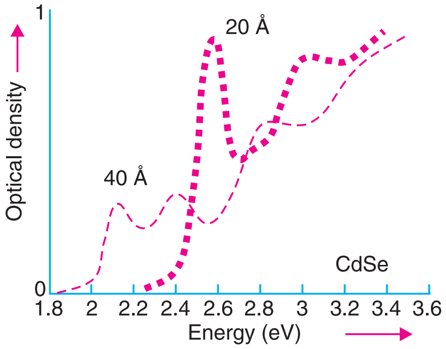


Fig. 79.10

There is a significant shift in the optical absorption spectra toward the blue as the particle size is reduced. There are two situations, called the *weak-confinement* and the *strong-confinement regimes*.

(*i*) In the weak regime the particle radius is larger than the radius of the electron-hole pair, but the range of motion of the exciton is limited, which causes a blue shift of the absorption spectrum.

(*ii*) When the radius of the particle is smaller than the orbital radius of the electron-hole pair, the motion of the electron and the hole become independent. The exciton does not exist. The hole and the electron have their own set of energy levels. So, new lines appear in the absorption spectrum, showing again a blue shift.

Fig. 79.10 shows the optical absorption spectrum of CdSe for two nanoparticles having sizes 20 Å and 40 Å, respectively measured at 10 K. The lowest energy absorption region, referred to as the *absorption edge*, is shifted to higher energy as the particle size decreases. Since the absorption edge is due to the band gap, this means that the band gap increases as particle size decreases. The intensity of the absorption increases as the particle size is reduced. The higher energy peaks are associated with the exciton. They shift to higher energies with the decrease in particle size. These effects are a result of the confinement of the exciton. As the particle size is reduced, the hole and the electron are forced closer together, and the separation between the energy levels changes.

79.4.2. Photofragmentation

Nanoparticles of Ge and Si undergo fragmentation when subjected to laser light from a Q-switched Nd:YAG laser. Some of the fissions are

$$Si_{12} + h\nu \rightarrow Si_6 + Si_6 \qquad ...(1)$$

$$Si_{20} + h\nu \rightarrow Si_{10} + Si_{10} \qquad ...(2)$$

Here, $h\nu$ is a photon of light energy.

The emerging products depend on the cluster size, the wavelength and the intensity of laser beam. When the cluster size is greater than 30 atoms, the fission occurs explosively.

79.4.3. Coulombic Explosion

Multiple ionizations of a cluster cause a rapid redistribution of the charges on the atoms of the cluster, making each atom more positive. When the Coulombic repulsion between the atoms in a cluster grows beyond the binding energy between the atoms, the cluster explodes and the atoms fly apart with great velocities. The phenomenon is called *Coulombic explosion*.

79.5 Rare Gas and Molecular Clusters

Nanoparticles can be assembled from rare gases such as krypton and xenon, and molecules such as water and carbon dioxide. Examples of such largest clusters are: $(Kr_{73})^{2+}$, $(Xe_{206})^{4+}$ and $(CO_2)^{4+}_{216}$.

1. Inert-Gas Clusters: Xenon clusters are formed by adiabatic expansion of a supersonic jet of the gas through a small capillary into a vacuum. The gas is then collected by a mass spectrometer, where it is ionized by an electron beam, and its mass: charge ratio measured. The most stable clusters occur at particles having 13, 19, 25, 55, 71, 87, and 147 atoms.

These numbers are called *structural magic numbers*. It means that clusters having a certain number of atoms are more stable than others. The forces that bond inert-gas atoms into clusters are weaker than those that bond metals and semiconducting atoms.

2. Superfluid Clusters: Clusters of ^{4}He and ^{3}He atoms formed by supersonic free-jet expansion of helium gas have been studied by mass spectrometry. Magic numbers are found at cluster sizes of N = 7,10,14,23,30 for ^{4}He, and N = 7,10,14,21,30 for ^{3}He. Superfluidity has been observed in He clusters having 64 and 128 atoms. In the state of *Bose-Einstein condensation* (the state where all the bosons occupy the lowest energy level), the liquid He^4 at 2.2 K becomes a superfluid. The viscosity of liquid He^4 drops to zero.

3. Molecular Clusters: Individual molecules can form clusters. One of the most common examples of this is the water molecule. At ambient conditions 80% of water molecules are bonded into clusters. Other examples of molecular clusters are $(NH_3)_n^+$, $(CO_2)_{44}$ and $(C_4H_8)_{30}$.

CARBON NANOSTRUCTURES

79.6 Carbon Clusters

79.6.1. Small Carbon Clusters

Laser evaporation of a carbon substrate using the apparatus shown in Fig. 79.1 in a pulse of He gas is used to make carbon clusters. The neutral cluster beam is photoionized by a UV laser and analyzed by a mass spectrometer. Fig. 79.11 shows the mass spectrum of carbon clusters.

For the number of atoms *N* less than 30, there are clusters of every *N*, although some are more prominent than others. Calculations of the structure of small clusters by molecular orbital theory show the clusters have linear or closed nonplanar monocyclic geometries.

(*i*) The linear clusters have *sp* hybridization with carbon bond angle as 180°. The linear structures occur when *N* is odd. The open structures with 3, 11, 15, 19, and 23 carbons are more prominent and more stable.

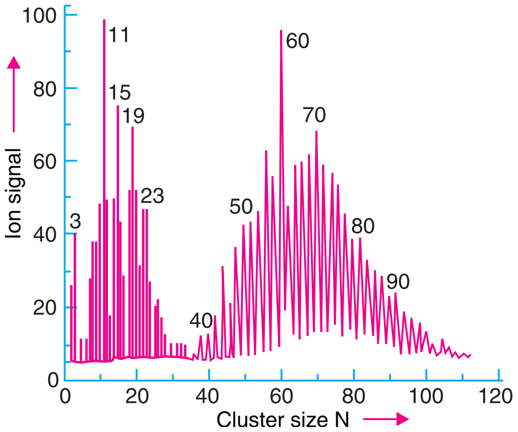


Fig. 79.11

The two other forms of hybridization occurring in carbon molecules are sp^2 and sp^3, with carbon bond angles measuring 120° and 109°28’ respectively.

(*ii*) Closed structures form when *N* is even. The closed structures have angles between the carbon bonds that differ from those predicted by the conventional hybridization concept. One such cluster composed of 60 atoms shows the highest intensity in the mass spectrum and hence the most stable.

79.6.2. Discovery of C_{60}

The evidence for the existence of a molecule containing 60 carbon atoms was first found while measuring the intensity of light coming from stars at different wavelengths. A pronounced reduction in the intensity of light at 220 nm falling on Earth was detected. This is referred to as *optical extinction*. It was attributed to light scattered from small particles of graphite that were believed to be present in the regions between the stars. This explanation for the optical extinction in the 220-nm region has been widely accepted by astronomers.

The IR studies involving the use of graphite arcs were carried out for further clarification and confirmation. Huffman and Kratschmer observed four IR lines that correspond closely to the lines of a molecule C_{60}, predicted theoretically to exist several years earlier. Other investigations using techniques including the mass spectrometry were then made that led to the confirmation of the existence of C_{60} molecule in 1990.

79.6.3. Structure of C_{60} and its Crystal

The C_{60} molecule has been named fullerene.

Fig. 79.12 shows the structure of the C_{60} *fullerene* molecule. It has 12 pentagonal (5 sided) and 20 hexagonal (6 sided) faces symmetrically arrayed to form a molecular ball. A soccer ball has the same geometric configuration as fullerene. These ball-like molecules bind with each other in the solid state to form a crystal lattice having a face centered cubic structure (Fig. 79.13).

In the lattice each C_{60} molecule is separated from its nearest neighbor by 1 nm. The molecules are held together by weak forces called van der Waals forces. Because C_{60} is soluble in benzene, single crystals of it can be grown by slow evaporation from benzene solutions.

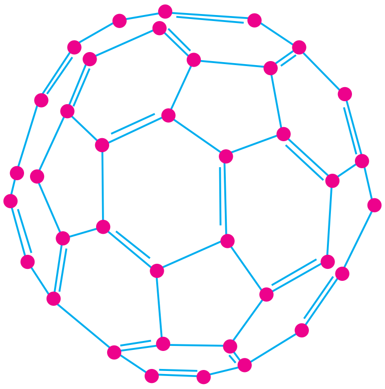


Fig. 79.12

79.6.4. Alkali-Doped C_{60}

In the FCC fullerene structure, 26% of the volume of the unit cell is empty. So alkali atoms can easily fit into the empty spaces between the molecular balls of the material. Fig. 79.13 shows the crystal lattice unit cell of C_{60} molecules (large spheres) doped with alkali atoms (dark circles).

The C_{60} crystal is an insulator, but when doped with an alkali atom it becomes electrically conducting.

79.6.5. Superconductivity in C_{60}

Some fullerenes have been derived from C_{60} by crystallizing C_{60} along with alkali metals whose atoms are placed in gaps between the C_{60} spheres. These are found to show superconductivity. The transition temperatures of materials of this class range from 15 K to about 48 K. The transition temperature T_c of the fullerene Cs_2RbC_{60} is 33 K.

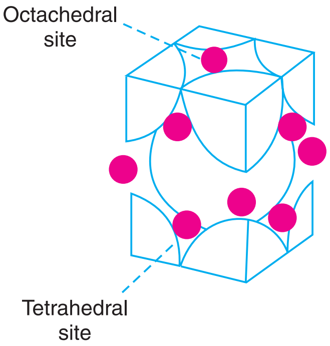


Fig. 79.13

79.7 Carbon Nanotubes

79.7.1. Fabrication of Carbon Nanotubes

There are three techniques of synthesizing carbon nanotubes: laser evaporation, carbon arc and chemical vapor deposition method.

1. Laser Evaporation Method

Fig. 79.14 shows the experimental arrangement for synthesizing carbon nanotubes by laser evaporation.

A quartz tube contains argon gas and a graphite target. The tube is centred inside a horizontal furnace maintained at 1200°C with portions toward the ends falling outside the furnace. A water-cooled copper collector is mounted within the tube at one of its ends, lying outside the furnace. The graphite target contains small amounts of cobalt and nickel that act as catalytic nucleation sites for the formation of the nanotubes.

An intense pulsed laser beam is incident on the target, evaporating carbon from the graphite. The argon then sweeps the carbon atoms from the high-temperature zone to the colder copper collector on which they condense into nanotubes. Tubes 10-20 nm in diameter and 100 μm long can be made by this method. This method produces tubes with closed ends.

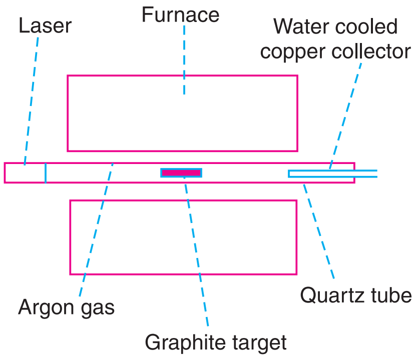


Fig. 79.14

2. Carbon Arc Method

A potential of 20-25V is applied across carbon electrodes of 5-20 μm diameter and separated by 1 mm at 500 torr pressure of flowing helium. Carbon atoms are ejected from the positive electrode and form nanotubes on the negative electrode. As the tubes form, the length of the positive electrode decreases, and a carbon deposit forms on the negative electrode.

This method produces tubes with closed ends.

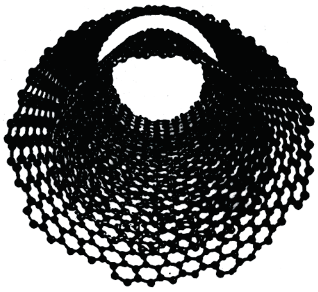


Fig. 79.15

(*i*) SWNT: To produce single-walled nanotubes, a small amount of cobalt, nickel, or iron is incorporated as a catalyst in the central region of the positive electrode. This method can produce single-walled nanotubes of diameters 1–5 nm with a length of 1 μm.

(*ii*) MWNT: If no catalysts are used, multi-walled nanotubes are produced. Fig. 79.15 shows a MWNT in which one tube is inside the another.

3. Chemical Vapor Deposition Method

This method involves decomposing a hydrocarbon gas such as methane (CH_4) at 1100°C. As the gas decomposes, carbon atoms are produced. Carbon atoms then condense on a cooler substrate that contains various catalysts such as iron. This method produced tubes with open ends.

79.7.2. Structure of Carbon Nanotubes

A nanotube can be formed when a graphite sheet is rolled up about the axis T (Fig. 79.16). $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ are the basis vectors of the two-dimensional unit cell. The C_h vector is called the *circumferential vector.* Three possible structures of carbon nanotubes constructed by rolling the graphite sheet about the T vector having different orientations in the graphite sheet are shown in Fig. 79.17.

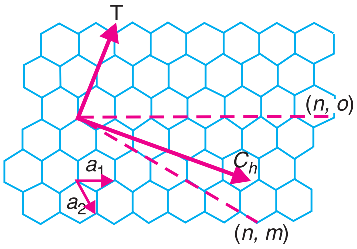
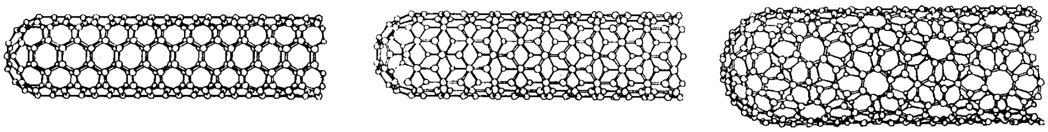


Fig. 79.16



(a) Armchair structure (b) Zigzag structure (c) Chiral structure

Fig. 79.17

(*i*) When the axis vector *T* is parallel to the C—C bonds of the carbon hexagons, the "*armchair*" structure is obtained.

(*ii*) The *zigzag* and the *chiral structures* are formed by rolling about a *T* vector having different orientations in the graphite plane, but not parallel to C—C bonds.

79.7.3. Electrical Properties of Carbon Nanotubes

1. Carbon nanotubes are metallic or semiconducting, depending on the diameter and chirality of the tube. *Chirality* refers to how the tubes are rolled with respect to the direction of the axis vector *T*. Synthesis generally results in a mixture of tubes two-thirds of which are semiconducting and one-third metallic. The metallic nanotubes have the armchair structure.

Fig. 79.18 is a plot of the energy bandgap of semiconducting chiral carbon nanotubes versus the reciprocal of the diameter. As the diameter of the tube increases, the bandgap decreases.

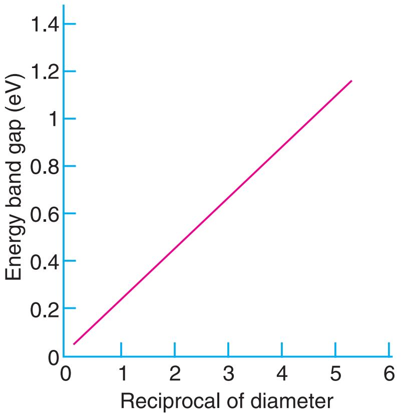


Fig. 79.18

2. The electronic structure of carbon nanotubes is studied using scanning tunnelling microscopy (STM). The data show that the electronic states of the tubes do not form a single wide electronic energy band, but instead split into one-dimensional subbands. These states can be modeled by a potential well having a depth equal to the length of the nanotube.

3. The voltage-current characteristics of a single metallic nanotube placed across two metal electrodes show step-like features. This resembles a field effect transistor (FET) constructed from a carbon nanotube.

4. In the metallic state the conductivity of the nanotubes is very high. They can carry extremely high current with density 10^9 A/cm^2. Copper wire fails at 10^6 A/cm^2 because resistive heating melts the wire. Carbon nanotubes have very few defects to scatter electrons, and thus a very low resistance. This is the reason for the high conductivity of the carbon nanotubes. High currents do not heat the carbon nanotubes in the same way that they heat copper wires.

5. Carbon nanotubes are very good conductors of heat, the thermal conductivity being almost twice as large as that of diamond.

6. *Magnetoresistance* is a phenomenon whereby the resistance of a material is changed by the application of a DC magnetic field. Carbon nanotubes display magnetoresistive effects at low temperature.

Fig. 79.19 shows the effect of a DC magnetic field on the resistance of nanotubes at the temperatures of 0.35 and 2.3 K. ΔR is the change in resistance of nanotubes at 2.3 and 0.35 K compared to their resistance *R* in zero magnetic field. This is a negative magnetoresistance effect because the resistance decreases with increasing DC magnetic field.

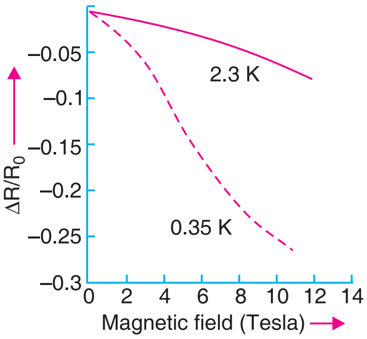


Fig. 79.19

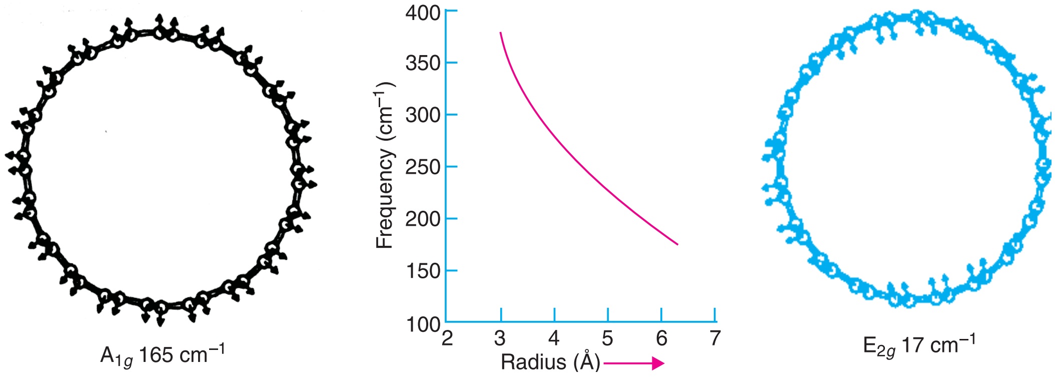
So the conductance $G = 1/R$, increases. This occurs because when a DC magnetic field is applied to the nanotubes, the conduction electrons acquire new energy levels associated with their spiraling motion about the field. For nanotubes, these levels, called *Landau levels*, lie very close to the topmost filled energy levels (the Fermi level). Thus there are more available states for the electrons to increase their energy, and the material is more conducting.

79.7.4. Vibrational Properties of Carbon Nanotubes

Fig. 79.20 (*a*) and (*c*) show two normal modes of vibration of carbon nanotubes.

(*i*) A_{1g} Mode: A_{lg} mode involves an 'in and out' oscillations of the diameter of the tube.

(*ii*) E_{2g} Mode: The E_{2g} mode involves a squashing of the tube where it squeezes down in one direction and expands in the perpendicular direction essentially oscillating between a sphere and an ellipse.



(*a*) (*b*) (*c*)

Fig. 79.20

The frequencies of these two modes are Raman-active and depend on the radius of the tube.

Fig. 79.20 (*b*) is a plot of the frequency of the Raman A_{lg} vibrational normal mode versus the radius of the nanotube.

It is now a standard practice to exploit this property for measuring the radius of carbon nanotubes.

79.7.5. Mechanical Properties of Carbon Nanotubes

(*i*) Carbon nanotubes are very strong.

(*ii*) The Young's modulus of carbon nanotubes ranges from 1.28 to 1.8 TPa. Young's modulus of steel is 0.21 TPa, which means that Young's modulus of carbon nanotubes is almost 10 times that of steel.

(*iii*) When carbon nanotubes are bent, they are very resilient. They buckle like straws but do not break, and can be straightened back without any damage. Carbon nanotubes have a very low density of defects in the structure of their walls. So they do not fracture when bent severely.

(*iv*) The tensile strength of carbon nanotubes is about 45 billion pascals. High-strength steel alloys break at about 2 billion pascals. Thus carbon nanotubes are about 20 times stronger than steel.

(*v*) Multi-walled nanotubes also have improved mechanical properties, but they are not as good as their single-walled counterparts. For example, multi-walled nanotubes of 200 nm diameter have a tensile strength of 0.007 TPa and a modulus of 0.6 TPa.

79.8 Applications of Carbon Nanotubes

1. Field Emission and Shielding

(*i*) When a small electric field is applied parallel to the axis of a nanotube, electrons are emitted at a very high rate from the ends of the tube. This effect, called *field emission,* has contributed to the development of flat panel displays.

(*ii*) The high electrical conductivity of carbon nanotubes means that they will be poor transmitters of electromagnetic energy. A plastic composite of carbon nanotubes could provide lightweight shielding material for electromagnetic radiation.

2. Computers: Carbon nanotubes with diameters of 2 nm have extremely low resistance. They can carry large currents without heating. So they could be used as interconnects.

Their very high thermal conductivity means that they can also serve as heat sinks, allowing heat to be rapidly transferred away from the chip.

3. Chemical Sensors: A field-effect transistor made of the chiral semiconducting carbon nanotubes is a sensitive detector of various gases.

4. Catalysis: Carbon nanotubes serve as catalysts for some chemical reactions.

5. Mechanical Reinforcement: Carbon nanotubes are used to increase the strength of plastic composites.

QUANTUM WELLS, WIRES, AND DOTS

79.9 Quantum Nanostructures

When the size or dimension of a material is continuously reduced from a large or macroscopic size, such as a meter or a centimeter, to a very small size, the properties remain the same at first, then small changes begin to occur, until finally when the size drops below 100 nm, dramatic changes in properties can occur. Three limiting classes of nanostructures emerge on the basis whether only one, only two or all three dimensions fall in nanometre scale.

1. Quantum Well: If one dimension is reduced to the nanometre range while the other two dimensions remain large, then we get a structure called *quantum well.*

2. Quantum Wire: If two dimensions are reduced to the nanometre range and one remains large, the structure is referred to as a *quantum wire.*

3. Quantum Dot: The extreme case of this process of size reduction in which all three dimensions reach the low nanometer range is called a *quantum dot.*

The word *quantum* is associated with these three types of nanostructures because the changes in properties arise from the quantum-mechanical nature of physics in the domain of the ultrasmall.

Fig. 79.21 (*a*) shows formation of rectangular nanostructures by progressive reduction in size.

Fig. 79.21 (*b*) shows formation of curvilinear nanostructures by progressive reduction in size.

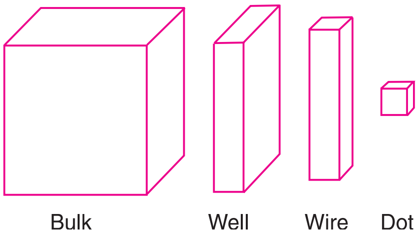


Fig. 79.21 (a)

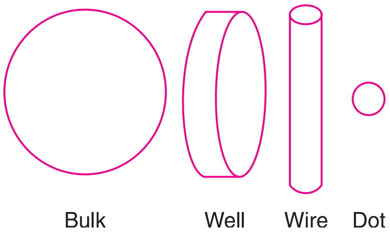


Fig. 79.21 (b)

79.10 Preparation of Quantum Nanostructures

There are two approaches to the preparation of a nanostructure, called (*i*) *bottom-up* approach, and (*ii*) *top-down* approach.

(*i*) Bottom-up method: The procedure is to collect, consolidate, and fashion individual atoms and molecules into the structure. This is carried out by a sequence of chemical reactions controlled by catalysts.

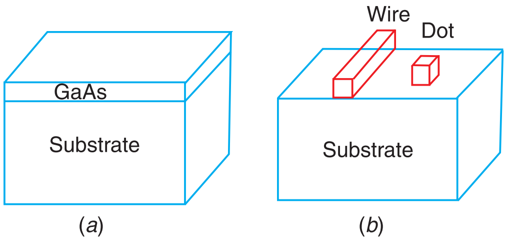
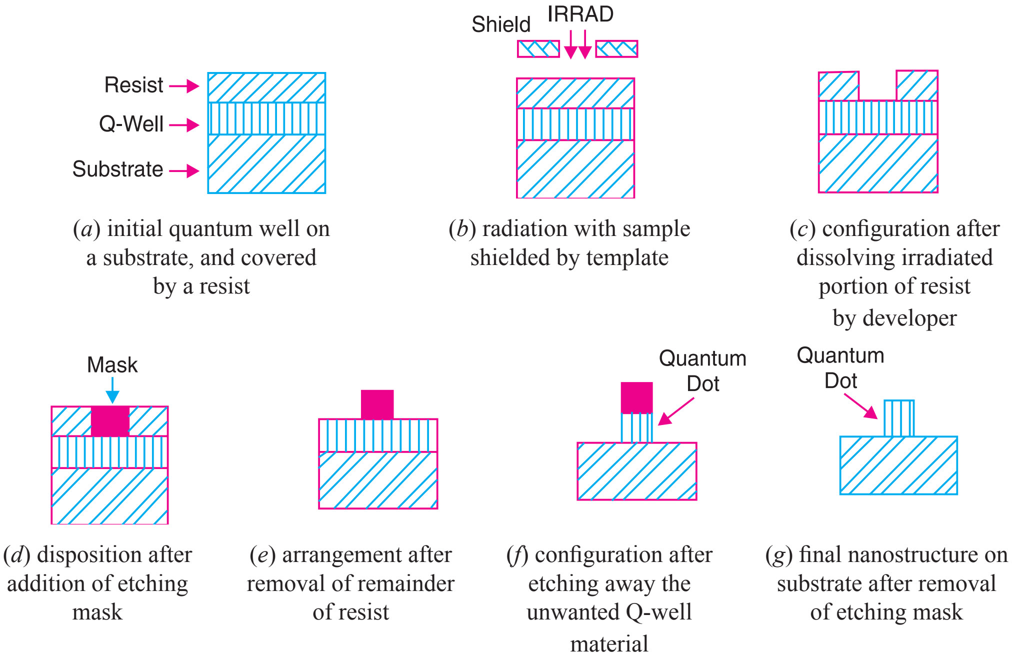


Fig. 79.22

(*ii*) Top-down method: The *top-down method* starts with a large-scale object or pattern and gradually reduces its dimension or dimensions. This is accomplished by a technique called *lithography* which shines radiation through a template on to a surface coated with a radiation-sensitive resist. The resist is then removed and the surface is chemically treated to produce the nanostructure. A typical resist material is the polymer polymethyl methacrylate $[C_5O_2H_8]_n$.

The lithographic process is illustrated by starting with a Gallium arsenide (GaAs) square quantum well located on a substrate [Fig. 79.22 (*a*)]. The final product to be produced from the material (GaAs) of the quantum well is either a quantum wire or a quantum dot [Fig. 79.22 (*b*)].

Fig. 79.23 shows the steps in the formation of a quantum wire or quantum dot by *electron-beam lithography*.



(*a*) initial quantum well on a substrate, and covered by a resist

(*b*) radiation with sample shielded by template

(*c*) configuration after dissolving irradiated portion of resist by developer

(*d*) disposition after addition of etching mask

(*e*) arrangement after removal of remainder of resist

(*f*) configuration after etching away the unwanted Q-well material

(*g*) final nanostructure on substrate after removal of etching mask

Fig. 79.23

(*a*) The sample quantum well is coated with a radiation-sensitive resist.

(*b*) The sample is then irradiated by an electron beam in the region where the nanostructure will be located. The radiation chemically modifies the exposed area of the resist so that it becomes soluble in a developer.

(*c*) The irradiated portions of the resist are removed, using the developer.

(*d*) An etching mask is inserted into the hole or groove in the resist.

(*e*) The remaining parts of the resist are now taken out.

(*f*) The areas of the quantum well not covered by the etching mask are chemically etched away to produce the quantum structure shown in Fig. 79.23 (*f*) covered by the etching mask.

(*g*) Finally the etching mask is removed, if necessary, to provide the desired quantum structure. The resulting structure may be a quantum wire or a quantum dot, depending on the shape and the size of the groove.

Quantum-Dot Array: The lithographic technique is used to make more complex quantum structures than the quantum wire and quantum dot. Starting with a multiple quantum-well structure, we can produce a quantum dot array. A quantum-dot array produces a greatly enhanced photoluminescent output of light.

79.11 Size and Dimensionality Effects

1. Size Effects: The genesis of the remarkably different properties of quantum nanostructures lies in their drastically reduced dimensions. As the size on nanoscale reduces, the fraction of atoms on the surface of the nanocrystal grows larger and larger. Consequently, the number of conduction electrons in metals and semiconductors drops considerably and the electron mean free path becomes size-dependent. The resistivity of a polycrystalline conductor or semiconductor composed of microcrystallites with diameters significantly greater than the mean free path resembles that of a network of interconnected resistors. But when the microcrystallite dimensions approach or become less than the mean free path, the resistivity depends mainly on scattering off boundaries between crystallites. Both types of metallic nanostructures are common.

Another size effect arises from the level of doping of a semiconductor. For typical doping levels of 10^{14} to 10^{18} donors/cm^3 a quantum-dot cube of side 100nm would have, on the average, from 10^{-1} to 10^3 conduction electrons. A smaller quantum-dot cube of side 10 nm would have, on the average, 10^{-4} conduction electrons for the 10^{14} doping level and 1 electron for the 10^{18} doping level. Similar estimates have been made for quantum wells and quantum wires.

The following Table gives conduction electron content of smaller size (on left) and larger size (on right) quantum structures containing donor concentrations of $10^{14} – 10^{18}$cm^{-3}.

Quantum Structure	*Size*	*Electron Content*	*Size*	*Electron Content*
Bulk material	—	$10^{14} – 10^{18}$ cm^{-3}	—	$10^{14} – 10^{18}$ cm^{-3}
Quantum well	10 nm thick	$1 – 10^4$ μm^{-2}	100 nm thick	$10 – 10^5$ μm^{-2}
Quantum wire	10 × 10–nm cross section	$10^{-2} – 10^2$ μm^{-1}	100 nm × 100 nm cross section	$1 – 10^4$ μm^{-1}
Quantum dot	10 nm on a side	$10^{-4} – 1$	100 nm on a side	$10^{-1} – 10^3$

2. Conduction Electrons and Dimensionality

Quantum Structure	**Delocalization Dimensions**	**Confinement Dimensions**
Bulk conductor	3 (x, y, z)	0
Quantum well	2(x, y)	1 (z)
Quantum wire	1(z)	2(x, y)
Quantum dot	0	3(x, y, z)

(*i*) In a copper wire (bulk conductor) the conduction electrons are delocalized and move freely throughout the entire conducting wire. All the wire dimensions are very large compared to the distances between atoms.

The situation changes when one or more dimensions of the copper becomes so small that it approaches several times the spacings between the atoms in the lattice. When this occurs, the delocalization is impeded, and the electrons experience confinement.

(*ii*) Consider a flat plate of copper that is 10 cm long, 10 cm wide, and only 3.6 nm thick. This thickness corresponds to the length of only 10 unit cells. This means that 20% of the atoms are in unit cells at the surface of the copper. The conduction electrons would be delocalized in the plane of the plate, but confined in the narrow dimension. This configuration is called a *quantum well*.

(*iii*) A *quantum wire* is a structure such as a copper wire that is long in one dimension, but has a nanometer size as its diameter. The electrons are delocalized and move freely along the wire, but are confined in the transverse directions.

(*iv*) A *quantum dot* may have the shape of a tiny cube, a short cylinder, or a sphere with low nanometer dimensions. It exhibits confinement in all three spatial dimensions. So there is no delocalization.

The above four cases are summarized in the Table.

3. Partial Confinement

The conduction electrons in nanostructures can be partially confined and partially delocalized, depending on the shape and the dimensions of .the structure.

One limiting case is a *quantum dot* in which they are totally confined.

The other limiting case is a *bulk material*, in which they are all delocalized.

The intermediate cases are a quantum wire and a quantum well which involve partial confinement.

A *quantum wire* is long in one dimension and very small in its transverse directions. The quantum wire exhibits electron confinement in two dimensions and delocalization in one dimension.

A *quantum well* is a flat plate nanosized in thickness and much larger in length and width. The quantum well exhibits electron confinement in one dimension and delocalization in two dimensions.

The following Table lists the number of electrons $N(E)$ and density of states $D(E) = dN(E)/dE$ as a function of the energy E for these four cases.

			Dimensions	
Type	***Number of electrons, N(E)***	***Density of states, D(E)***	***Delocalized***	***Confined***
Q-dot	$N(E) = K_0 \sum d_i \Theta(E - E_{iw})$	$D(E) = K_0 \sum d_i \delta(E - E_{iw})^2$	0	3
Q-wire	$N(E) = K_1 \sum d_i (E - E_{iw})^{1/2}$	$D(E) = \frac{1}{2} K_1 \sum d_i (E - E_{iw})^{-1/2}$	1	2
Q-well	$N(E) = K_2 \sum d_i (E - E_{iw})$	$D(E) = K_2 \sum d_i$	2	1
Bulk				
Conductor	$N(E) = K_3 (E)^{3/2}$	$D(E) = \frac{3}{2} K_3 (E)^{1/2}$	3	0

The summations in these expressions are over the various confinement well levels i.

K_0, K_1, K_2 and K_3 are constants.

The degeneracies d_i refer to potential well energy levels.

E_{iw} refers to the energy of the level i in the potential well.

The Heaviside step function $\Theta(x) = 0$ for $x < 0$ and 1 for $x > 0$.

The delta function $\delta(x) = 0$ for $x \neq 0$, and ∞ for $x = 0$, and integrates to a unit area.

Energy Dependence Curves of Number of Electrons and Density of States

Fig. 79.24 shows number of electrons $N(E)$ (left side) and density of states $D(E)$ (right side) plotted against the energy E for three quantum structures and their bulk counterpart in the square well-Fermi gas approximations.

(*i*) The number of electrons $N(E)$ increases with the energy E for all the four structures. So the four types vary only qualitatively from each other.

(*ii*) The density of states $D(E)$ differ dramatically for each of the three nanostructure types. This means that the nature of the dimensionality and of the confinement associated with a particular nanostructure have a pronounced effect on its properties. These considerations can be used to predict properties of nanostructures. We can also identify types of nanostructures from their properties.

Some of the properties of solids that depend on the density of states are: electron heat capacity, Pauli susceptibility, the intensity of x-ray emission, electron and hole concentrations in semiconductors, the superconducting energy gap, and Josephson junction tunneling in superconductors. Experimental measurements of these various properties permit us to determine the form of the density of states $D(E)$, both at the Fermi level E_F and over a broad range of temperature.

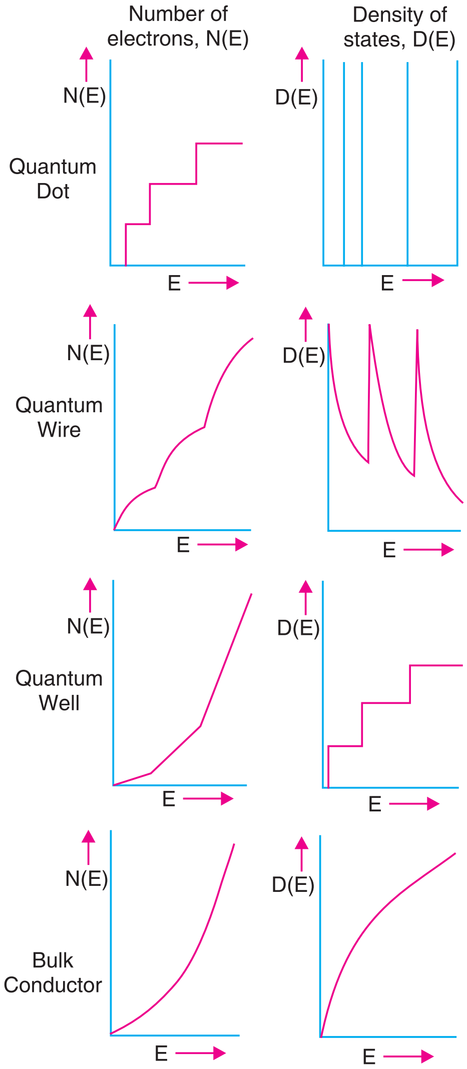


Fig. 79.24

79.12 Applications of Quantum Nanostructures

1. Infrared Detectors

Infrared transitions involving energy levels of quantum wells are used for the operation of infrared photodetectors. The infrared detectors depend on the presence of discrete energy levels in a quantum well between which transitions in the infrared spectral region can be induced.

2. Quantum Dot Laser

Fig. 79.25 shows the structure of a quantum dot near-infrared laser fabricated by Park *et al*.

In the quantum-dot laser, the quantum dots play the role of the active atoms.

The quantum-dot laser diode is grown on an n-doped GaAs substrate (not shown). The top p-metal layer has a GaAs contact layer immediately below it. Between this contact layer above and the GaAs substrate (not shown) below the diagram, there are a pair of 2-μm-thick $A1_{0.85}Ga_{0.15}$ As cladding or bounding layers that surround a 190-nm-thick waveguide made of $A1_{0\cdot 05}Ga_{0.95}$ As. The waveguide plays the role of conducting the emitted light to the exit ports at the edges of the structure.

Details of the waveguide region

The inset at the bottom of the figure shows details of the 190-nm-wide ($Al_{0.85}Ga_{0.15}As$ cladded) waveguide region. Centered in the waveguide (dark horizontal stripe on the figure labeled QD) is a 30-nm-thick GaAs region, and centered in this region are 12 monolayers of $In_{0.5}Ga_{0.5}As$ quantum dots with a density of $1.5 \times 10^{10}/cm^2$. The quantum dots do the lasing.

The cavity is typically 1 to 5 mm long and 4 to 60 μm wide. The faces of the laser are coated with $ZnSe/MgF_2$ high-reflectivity (> 95%) coatings. The coatings reflect light back and forth inside to augment the stimulated emission. The threshold current for CW operation at room temperature is 4.1-mA. The output signal has the wavelength of 1.32 μm. The laser light exits through the lateral edge of the laser.

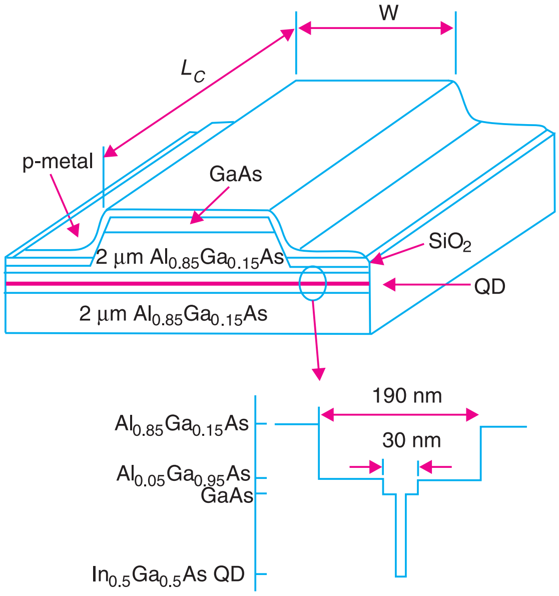


Fig. 79.25

NANOMATERIALS: FABRICATION BY PHYSICAL TECHNIQUES

79.13 Photolithography

Photolithography is the most commonly applied technique in the mass production of integrated circuits.

Fig. 79.26 outlines the basic steps in the photolithographic process of transferring images in the mask to the substrate surface.

- The resist material is applied as a thin coating over some base.
- Selected areas of the resist are then exposed to light through a mask in an image-wise fashion.
- The developer is now brought in contact with the exposed resist. A positive tone or a negative tone image of the mask is obtained after the developing process, depending on whether the exposed areas of the resist are more or less soluble in the developer as compared to the unexposed areas. The three-dimensional relief image thus produced in the resist material is a replica of the opaque and transparent areas of the mask.
- The areas of resist that remain following the imaging and developing processes are used to mask the underlying substrate for subsequent etching or other image transfer steps. The resist material resists the etchant and prevents it from attacking the underlying substrate in those areas where it remains in place after development.

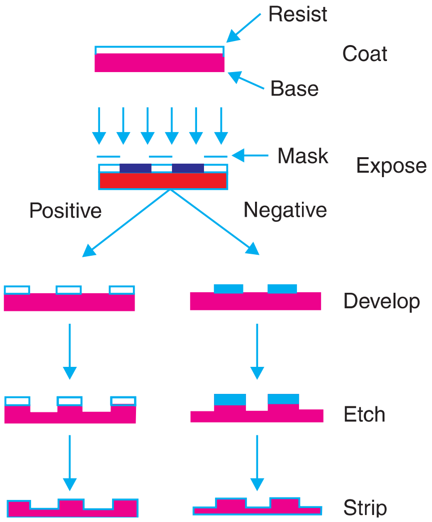


Fig. 79.26

- The resist is removed by stripping to produce a positive or negative tone relief image in the underlying substrate.

There are two basic photolithographic approaches:

(1) Shadow printing, which can be further divided into contact printing (or contact-mode printing) and proximity printing, and

(2) Projection printing.

The resolution or minimum size of elements

The *resolution* or minimum size of elements is set by diffraction effects (for photons).

Diffraction sets the limit of the maximum resolution or the minimum size of the individual elements by photolithography, which can be obtained.

- The theoretical resolution capability of shadow photolithography with a mask consisting of equal lines and spaces of width b is given by:

$$2b_{\min} = 3\sqrt{\lambda\left(s + \frac{d}{2}\right)} \qquad ...(1)$$

Here, $2b$ = grating period ($1/2b$ is the fundamental spatial frequency v)
s = gap width maintained between the mask and the photoresist surface
λ = wavelength of the exposing radiation and
d = photoresist thickness

For hard contact printing, $s = 0$, and from the equation, the maximum resolution for 400 nm wavelength light and a 1 μm thick resist film is slightly less than 1 μm.

- In contact-mode photolithography, the mask and wafer are in intimate contact. Thus this method can transfer a mask pattern into a photoresist with almost 100% accuracy and provides the highest resolution. Other photolithographic techniques can approach but not exceed its resolution capabilities. However, the maximum resolution is seldom achieved because of dust on substrates and non-uniformity of the thickness of the photoresist and the substrate. Such problems can be avoided in proximity printing.
- In proximity printing, a gap between the mask and the wafer is introduced. However, increasing the gap degrades the resolution by expanding the penumbral region caused by diffraction. The difficulties in proximity printing include the control of a small and very constant space between the mask and wafer, which can be achieved only with extremely flat wafers and masks.
- In projection printing, the mask and the wafer substrate are separated by several centimetres. Lens elements are used to focus the mask image onto the wafer substrate. Because of lens imperfections and diffraction considerations, projection techniques have lower resolution capability than that provided by shadow printing. The resolution limit in projection photolithography is determined by the Rayleigh's equation.

The minimum resolvable feature, R, and the corresponding depth of focus (DOF) are given by the following :

$$R = \frac{k_1\lambda}{NA} \qquad ...(2)$$

$$\text{DOF} = \frac{k_2\lambda}{(NA)^2} \qquad ...(3)$$

Here λ is the exposure wavelength, k_1 and k_2 are constants that depend on the specific resist material, process technology and image-formation techniques used.

NA is the numerical aperture of the optical system and is defined as

$$NA = n \sin\theta \qquad ...(4)$$

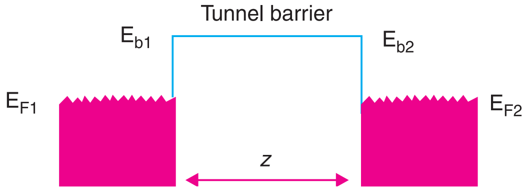
Here, n = index of refraction in image space and

θ = maximum cone angle of the exposure light beam.

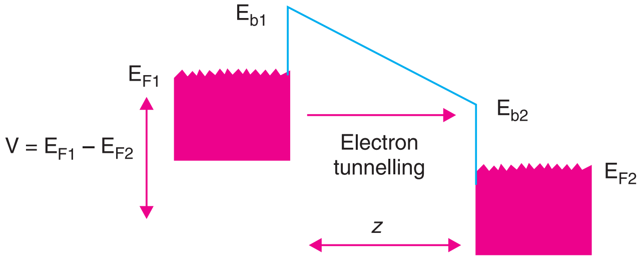
Conventional photolithography is capable of fabricating features of 200 nm and above.

79.14 Scanning Tunnelling Microscopy (STM)

Principle. STM relies on electron tunnelling, which is a quantum-mechanical phenomenon. Consider two flat surfaces of a metal or semiconductor, separated by an insulator or vacuum (Fig. 79.27).



(*a*) with no electric field applied between the solids.



(b) with an electric field applied between the solids.

Fig. 79.27

E_{F1} and E_{F2} are the Fermi levels of the respective solids.

Energies of the electrons in the solids are indicated by the shaded areas up to E_{F1} and E_{F2}.

The applied bias V is $E_{F1} - E_{F2}$.

z is the distance between the two solids or the thickness of the insulator.

(*a*) Electrons in the material cannot transfer from one surface to another through the insulator, since there is an energy barrier.

(*b*) When a voltage is imposed between the two surfaces, the shape of the energy barrier changes [see Fig. 79.27 (*b*)]. There is a driving force for electrons to move across the barrier by tunnelling. This results in a small current when the distance is sufficiently small so that the electron wave functions extended from the two surfaces overlap.

The tunnelling current, *I*, is given by:

$$\mathrm{I} \propto e^{-2kz} \qquad \text{...(1)}$$

k is given by:

$$k = \frac{\sqrt{2m\,(V - E)}}{h} \qquad \text{...(2)}$$

Here, *m* is the mass of an electron, *h* is Planck's constant, *E* is the energy of electron, and *V* is the potential in the insulator.

Block diagram of scanning tunnelling microscope (STM)

Fig. 79.28 shows the main components of STM.

A conducting wire having a very sharp tip is positioned above the surface of a sample. When the tip moves back and forth across the sample surface at very small intervals, the height of the tip is continually adjusted to keep the tunnelling current constant. The tip positions are used to construct a topographic map of the surface.

- The extremely sharp tip is usually made of metals or metal alloys, such as tungsten or PtIr alloy. The tip is mounted on to a three-dimensional positioning stage made of an array of piezoelectrics. The tip moves above the sample surface in three dimensions accurately controlled by the piezoelectric arrays. The distance between the tip and the sample surface falls between 0.2 and 0.6 nm. A tunneling current in the scale of 0.1 – 10nA is generated.
- The scanning resolution is about 0.01 nm in XY direction and 0.002 nm in Z direction. This results in atomic scale resolution in the three-dimensional image. The STM provides atomic-scale resolution of *surfaces*. Its inventors, G. Binnig and H. Rohrer of IBM Zurich, won the 1986 Nobel Prize for this work.

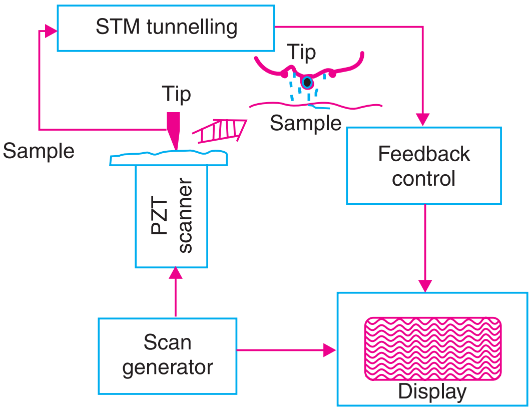


Fig. 79.28

The Tunnelling Current

The configuration of a STM is a tip-planar surface geometry.

The tunnelling current is given by:

$$I = C\rho_t\rho_s e^{-zk^{\frac{1}{2}}} \qquad \text{...(3)}$$

Here, z is the distance between the tip and the planar surface or sample,

ρ_t is the tip electronic structure,

ρ_s is the sample electronic structure, and

C is a constant dependent on the voltage applied between the tip and the sample surface.

- The tunnelling current decays exponentially with the tip-sample distance. For example, a 0.1 nm decrease in the distance will increase the tunneling current by one order of magnitude. This quantum mechanical property is used in the STM.

Operating modes. STM is operated in two modes.

(*i*) Constant current mode: In constant current imaging, a feed-back mechanism is enabled that a constant current is maintained while a constant bias is applied between the sample and tip. As the tip scans over the sample, the vertical position of the tip is altered to maintain the constant separation.

(*ii*) Constant height mode: In this imaging mode, constant height and bias are simultaneously maintained. A variation in current results as the tip scans the sample surface because a topographic structure varies the tip-sample separation.

- The constant current mode produces a contrast directly related to electron charge density profiles, whereas the constant height mode permits faster scan rates.
- A limitation of the STM is the requirement for a conducting sample. A variation on the STM for insulators is the atomic-force microscope (AFM) which operates on the atomic force between the sample atoms and the probe atoms (in contrast to the tunneling current for the STM).

79.15 Atomic Force Microscopy (AFM)

AFM is a modification of STM for dielectric materials. AFM monitors the force between the sample surface and the probe tip. A variety of tip-sample interactions may be measured by an AFM, depending on the separation.

(*i*) At short distances, the van der Waals interactions are predominant.

Van der Waals force consists of interactions of three components: permanent dipoles, induced dipoles and electronic polarization.

(*ii*) Long-range forces become significant when the tip-sample distance is large. Examples of such forces include electrostatic attraction or repulsion, current-induced or static-magnetic interactions, and capillary forces due to the condensation of water between the sample and tip.

Working of an atomic force microscope.

In AFM, the motion of a cantilever beam with an ultra small mass is measured. The force required to move this beam through measurable distance (10^{-5} nm) can be as small as 10^{-18} N.

Fig. 79.29 shows how the AFM works.

The instrument consists of a cantilever with a nanoscale tip, a laser pointing at the end of a cantilever, a mirror and a photodiode collecting the reflected laser beam, and a three-dimensional positioning sample stage which is made of an array of piezoelectrics.

The images are generated by scanning the tip across the surface. The AFM measures the minute upward and downward deflections of the tip cantilever while maintaining a constant force of contact.

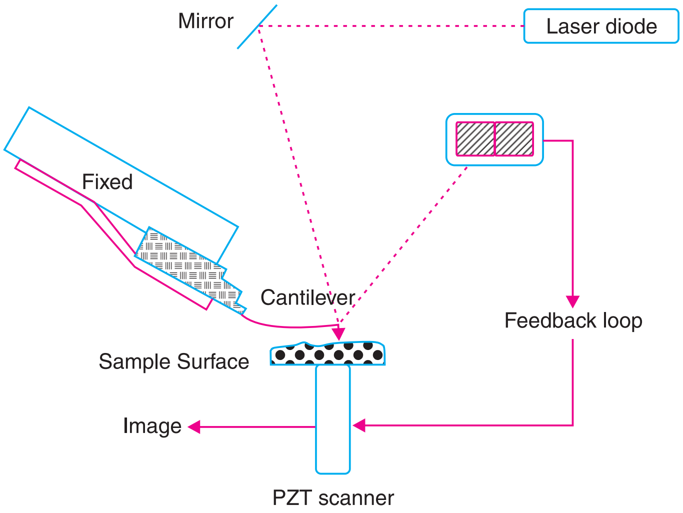


Fig. 79.29

The sample is mounted on a scanner. The cantilever and tip are positioned near the surface with a macroscopic positioning device. A photo diode records the position of a laser beam, reflected off the top of the cantilever.

AFM's can achieve a resolution of 10 pm.

Difference between STM and AFM:

STM: monitors the tunnelling current between the surface and the probe tip.

AFM: monitors the force between the surface and the probe tip.

EXERCISE

1. What are nanomaterials? Describe briefly different types of nanoparticles and discuss their structures.
2. Describe the structure and properties of various nanoparticles.
3. Discuss in brief the essentials of jellium model and its success.
4. Describe the structure and properties of carbon clusters.
5. What are fullerenes? Describe in detail the structure and properties of fullerite. Mention present and projected applications of fullerenes.
6. Describe in brief how carbon nanotubes are formed. Discuss their structure, characteristic properties and applications.
7. What are quantum nanostructures? Give a brief description of their fabrication.
8. Discuss the salient features of quantum nanostructures and describe their potential applications.
9. Discuss the basic principles of photolithography. Give experimental and theoretical bases of the technique.
10. Describe the principle and working of a scanning tunnelling microscope. Discuss briefly its application and the resolution achieved.
11. Describe the principle and working of an atomic force microscope. Give a comparative discussion on its application with reference to STM.